

FIGURE 1

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGGATTTACAAAAGGTGCAGGTATG
AGCAGGTCTGAAGACTAACATTTTGTGAAGTTGTAAAACAGAAAACCTGTTAGAAATGTGGTGGT
TTCAGCAAGGCCTCAGTTTCCTTCCTTCAGCCCTTGTAATTTGGACATCTGCTGCTTTCATATTT
TCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTTACCTTATATCAGTGACACTGG
TACAGTAGCTCCAGAAAAATGCTTATTTGGGGCAATGCTAAATATTGCGGCAGTTTTATGCAATTG
CTACCATTTATGTTTCGTTATAAGCAAGTTCATGCTCTGAGTCCTGAAGAGAACGTTATCATCAAA
TTAAACAAGGCTGGCCTTGIACTTGGAATACTGAGTTGTTTAGGACTTTCTATTGTGGCAAACCTT
CCAGAAAAACAACCCCTTTTGTGTCACATGTAAGTGGAGCTGTGCTTACCTTTGGTATGGGCTCAT
TATATATGTTTGTTCAGACCATCCTTTCCTACCAAATGCAGCCCAAAATCCATGGCAAACAAGTC
TTCTGGATCAGACTGTTGTTGGTTATCTGGTGTGGAGTAAGTGCACTTAGCATGCTGACTTGCTC
ATCAGTTTTGCACAGTGGCAATTTTGGGACTGATTTAGAACAGAACTCCATTGGAACCCCGAGG
ACAAAGGTTATGTGCTTCACATGATCACTACTGCAGCAGAATGGTCTATGTCATTTTCCTTCTTT
GGTTTTTTCCTGACTTACATTCGTGATTTTCAGAAAATTTCTTTACGGGTGGAAGCCAATTTACA
TGGATTAACCCCTCTATGACACTGCACCTTGCCCTATTAACAATGAACGAACACGGCTACTTTCCA
GAGATATTTGATGAAAGGATAAAATATTTCTGTAATGATTATGATTCTCAGGGATTGGGGAAAGG
TTCACAGAAGTTGCTTATTCTTCTCTGAAATTTCAACCACTTAATCAAGGCTGACAGTAACACT
GATGAATGCTGATAATCAGGAAACATGAAAGAAGCCATTTGATAGATTATCTAAAGGATATCAT
CAAGAAGACTATTAAAAACACCTATGCCTATACTTTTTTATCTCAGAAAATAAAGTCAAAAGACT
ATG

FIGURE 2

<subunit 1 of 1, 266 aa, 1 stop

<MW: 29766, pI: 8.39, NX(S/T): 0

MWWFQQGLSFLPSALVIWTSAAFIYSYITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNIAAV
LCIATIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSGAVLTFG
MGSLYMFVQTILSYQMOPKIHGKQVFWIRLLVIWCGVSALSMLTCSSVLHSGNFGTDLEQKLHW
NPEDKGYVLHMITTAAEWSMSFSFFGFFLTYYIRDFQKISLRVEANLHGLTLYDTAPCPINNERTR
LLSRDI

Important features:

Type II transmembrane domain:

amino acids 13-33

Other Transmembrane domains:

amino acids 54-73, 94-113, 160-180, 122-141

N-myristoylation sites.

amino acids 57-63, 95-101, 99-105, 124-130, 183-189

FIGURE 3

CGGACGCGTGGGCGGACGCGTGGGGGAGAGCCGAGTCCCGGCTGCAGCACCTGGGAGAAGGCAGACC
 GTGTGAGGGGGCCTGTGGCCCCAGCGTGCTGTGGCCTCGGGGAGTGGGAAGTGGAGGCAGGAGCCTTC
 CTTACACTTCGCCATGAGTTTCCTCATCGACTCCAGCATCATGATTACCTCCAGATACTATTTTTTG
 GATTTGGGTGGCTTTTCTTCATGCGCCAATTGTTTAAAGACTATGAGATACGTAGTATGTTGTACAG
 GTGATCTTCTCCGTGACGTTTGCAATTTCTTGCACCATGTTTGAGCTCATCATCTTTGAAATCTTAGG
 AGTATTGAATAGCAGCTCCCGTTATTTTCACTGGAAAATGAACCTGTGTGTAATCTGCTGATCCTGG
 TTTTCATGGTGCCTTTTACATTGGCTATTTTATTGTGAGCAATATCCGACTACTGCATAAACAACGA
 CTGCTTTTTTCTGTCTCTTATGGCTGACCTTTATGTATTTCTTCTGGAACTAGGAGATCCCTTTCC
 CATTCTCAGCCCAAACATGGGATCTTATCCATAGAACAGCTCATCAGCCGGGTGGTGTGATTGGAG
 TGACTCTCATGGCTCTTCTTTCTGGATTTGGTGTGTCAACTGCCCATACACTTACATGTCTTACTTC
 CTCAGGAATGTGACTGACACGGATATTCTAGCCCTGGAACGGCGACTGCTGCAAACCATGGATATGAT
 CATAAGCAAAAAGAAAAGGATGGCAATGGCACGGAGAACAATGTTCCAGAAGGGGGAAGTGCATAACA
 AACCATCAGGTTTCTGGGGAATGATAAAAAGTGTTACCACTTCAGCATCAGGAAGTGAATCTTACT
 CTTATTCAACAGGAAGTGGATGCTTTGGAAGAATTAAGCAGGCAGCTTTTTCTGGAAACAGCTGATCT
 ATATGCTACCAAGGAGAGAATAGAATACTCCAAAACCTTCAAGGGGAAATATTTTAAATTTCTTGGTT
 ACTTTTTCTCTATTTACTGTGTTTGGAAAATTTTCATGGCTACCATCAATATTGTTTTTGATCGAGTT
 GGGAAAACGGATCCTGTCACAAGAGGCATTGAGATCACTGTGAATTATCTGGGAATCCAATTTGATGT
 GAAGTTTGGTCCCAACACATTTCTTCTATTCTTGTGGAATAATCATCGTCACATCCATCAGAGGAT
 TGCTGATCACTCTTACCAAGTTCTTTTATGCCATCTCTAGCAGTAAGTCTCCAATGTCATTGTCTCTG
 CTATTAGCACAGATAATGGGCATGTACTTTGTCTCCTCTGTGCTGCTGATCCGAATGAGTATGCCTTT
 AGAATACCGCACCATAATCACTGAAGTCCTTGGAGAACTGCAGTTCAACTTCTATCACCGTTGGTTTG
 ATGTGATCTTCTGGTCAGCGCTCTCTCTAGCATACTCTTCTCTATTTGGCTCACAAACAGGCACCA
 GAGAAGCAAATGGCACCTTGAACTTAAGCCTACTACAGACTGTTAGAGGCCAGTGGTTTCAAAATTTA
 GATATAAGAGGGGGGAAAAATGGAACAGGGCCTGACATTTTATAAACAAACAAAATGCTATGGTAGC
 ATTTTTCACCTTCATAGCATACTCCTTCCCGTCAGGTGATACTATGACCATGAGTAGCATCAGCCAG
 AACATGAGAGGGAGAACTAACTCAAGACAATACTCAGCAGAGAGCATCCCGTGTGGATATGAGGCTGG
 TGTAGAGGCGGAGAGGAGCCAAGAACTAAAGGTGAAAAATACACTGGAACCTCTGGGGCAAGACATGT
 CTATGGTAGCTGAGCCAAACACGTAGGATTTCGTTTAAAGGTTACATGGAAGGTTATAGCTTTG
 CCTTGAGATTGACTCATTAAATCAGAGACTGTAACAAAAAAAAAAAAAAAAAGGGCGGCCGCG
 ACTCTAGAGTCGACCTGCAGAAGCTTGGCCGCCATGGCCCACTTGTATTGCAGCTTATAATG

$$= \frac{1}{\sqrt{2}} \begin{pmatrix} -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}$$

MSFLIDSSIMITSQILFFGFGWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCMTMFEIIFIIEILGV
LNSSSRFYHWMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLLFSCLLWLTFMYFFWKLGD
FPILSPKHGILSIEQLISRVGVIGVTLMALLSGGAVNCPYTYSYFLRNVTDTDILALERRLLQ
TMDMIISSKKRMAMARRTMFQKGEVHNKPSGFGWGMISVTTASGSENLTIIQQEVDAL EELS
RQLFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVFDRVGKTPVTRGIEI
TVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISSKSSNVIVLLLAQIMGY
FVSSVLLIRMSMPLEYRTIITEVLGELQNFYHRWFDVIFLVSALSSILFLYLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 5

AGCAGGGAAATCCGGATGTCTCGGTTATGAAGTGGAGCAGTGAGTGTGAGCCTCAACATAGTTCC
 AGAATCTCCATCCGACTAGTTATTTAGCATCTGCCTCTCATATCACCAGTGGCCATCTGAGGT
 GTTTCCTGGCTCTGAAGGGGTAGGCACGATGGCCAGGTGCTTCAGCCTGGTGTGCTTCTCACT
 TCCATCTGGACCACGAGGCTCCTGGTCCAAGGCTCTTTGCGTGCAGAAGAGCTTTCATCCAGGT
 GTCATGCAGAATTATGGGGATCACCTTGTGAGCAAAAAGGCGAACCAGCAGCTGAATTTACAG
 AAGCTAAGGAGGCTGTAGGCTGCTGGGACTAAGTTTGGCCGGCAAGGACCAAGTTGAAACAGCC
 TTGAAAGCTAGCTTTGAAACTTGCAGCTATGGCTGGGTGGAGATGGATTCTGGTTCATCTCTAG
 GATTAGCCCAAACCCCAAGTGTGGGAAAAATGGGTGGGTGTCTGATTGGAAGGTTCCAGTGA
 GCCGACAGTTTGCAGCCTATTGTTACAACCTCATCTGATACTTGGACTAATCGTGCATTCCAGAA
 ATTATCACCACCAAGATCCCATTCAACACTCAAACCTGCAACACAAACACAGAATTTATTTGT
 CAGTGACAGTACCTACTCGGTGGCATCCCCTTACTCTACAATACCTGCCCCCTACTACTCTCCTC
 CTGCTCCAGCTTCCACTTCTATTCCACGGAGAAAAAATTGATTGTGTCACAGAAGTTTTATG
 GAACTAGCACCATGTCTACAGAACTGAACATTGTTGAAAATAAAGCAGCATTCAAGAATGA
 AGCTGCTGGGTTTGGAGGTGTCCCCACGGCTCTGCTAGTGCTTGCTCTCCTCTCTTTGGTGTG
 CAGCTGGTCTTGGATTTTGTATGTCAAAGGTATGTGAAGGCTTCCCTTTTACAAACAAGAAT
 CAGCAGAAGGAAATGATCGAAACCAAAGTAGTAAAGGAGGAGAGGCAATGATAGCAACCCATA
 TGAGGAATCAAAGAAAACGATAAAAACCCAGAAGAGTCCAAGAGTCCAAGCAAACTACCGTGC
 GATGCCTGGAAGCTGAAGTTTAGATGAGACAGAAATGAGGAGACACACCTGAGGCTGGTTTCTTT
 CATGCTCCTTACCCTGCCCCAGCTGGGGAAATCAAAGGGCCAAAGAACCAGAAAGAAAGTCCA
 CCCTTGGTTCCCTAATGGAATCAGCTCAGGACTGCCATTGGACTATGGAGTGCACCAAGAGAAT
 GCCCTTCTCCTTATGTAACCCCTGTCTGGATCCTATCCTCCTACCTCCAAAGCTTCCCACGGCT
 TTCTAGCCTGGCTATGTCTAATAATATCCCACTGGGAGAAAGGAGTTTTCGAAAGTGCAGGAC
 CTAAAACATCTCATCAGTATCCAGTGGTAAAAAGGCTCCTGGCTGTCTGAGGCTAGGTGGGTG
 AAAGCCAAGGAGTCACTGAGACCAAGGCTTTCTCTACTGATTCCGCAGCTCAGACCCTTTCTTCA
 GCTCTGAAAGAGAAACACGTATCCACCTGACATGTCTTCTGAGCCCGGTAAAGAGCAAAAGAAAT
 GGCAGAAAAGTTTAGCCCCGTGAAAGCCATGGAGATTCTCATAACTTGAGACCTAATCTCTGTAAA
 GCTAAAATAAAGAAATAGAACAAGGCTGAGGATACGACAGTACACTGTGACAGGGACTGTAAAC
 ACAGACAGGGTCAAAGTGTCTTCTCTGAACACATTGAGTTGGAATCACTGTTTAGAACACACACA
 CTTACTTTTCTGGTCTCTACCACTGCTGATATTTCTCTAGGAAATATACTTTTACAAGTAACA
 AAAATAAAAACCTCTTATAAATTTCTATTTTATCTGAGTTACAGAAATGATTACTAAGGAAGATT
 ACTCAGTAATTTGTTTAAAAAGTAATAAAATTCACAAACATTGCTGAATAGCTACTATATGTC
 AAGTGCTGTGCAAGGTATTACACTCTGTAATTGAATATTATTCCTCAAAAAATGACATAGTAG
 AACGCTATCTGGGAAGCTATTTTTTTCAGTTTTGATATTTCTAGCTTATCTACTTCCAACTAAT
 TTTTATTTTTGCTGAGACTAATCTTATTCATTTTCTCTAATATGGCAACCATTTATAACCTTAATT
 TATTATTAACATACCTAAGAAGTACATTGTTACCTCTATATACCAAGCACATTTTAAAGTGCC
 ATTAACAAATGTATCACTAGCCCTCCTTTTTTCCAACAAGAAGGACTGAGAGATGCAGAAATATT
 TGTGACAAAAAATTAAAGCATTTAGAAACTT

FIGURE 7

CGCCGCGCTCCCCGACCCGCGGCCCCGCCACCGCGCGCTCCCGCATCTGCACCCGCGAGCCGGC
GGCCTCCCGGGCGGAGCGAGCAGATCCAGTCCGGCCCCGAGCGCAACTCGGTCCAGTCGGGGCGG
CGGCTGCGGGCGCAGAGCGGAGATCGAGCGGCTTGGGGCCACCCTGCTGTGCCTGCTGCTGGCGG
CGGCGGTCCCCACGGCCCCCGCGCCCGCTCCGACGCGACCTCGGCTCCAGTCAAGCCCCGGCCG
GCTCTCAGCTACCCGCGAGGAGGAGGCCACCCTCAATGAGATGTTCCGCGAGGTTGAGGAAGTAT
GGAGGACACGCAGCACAAATTGCGCAGCGCGGTGGAAGAGATGGAGGCAGAAGAAGCTGCTGCTA
AAGCATCATCAGAAGTGAACCTGGCAAACCTTACCTCCAGCTATCACAATGAGACCAACACAGAC
ACGAAGGTTGGAAATAATACCATCCATGTGCACCGAGAAATTCACAAGATAACCAACAACAGAC
TGGACAAATGGTCTTTTCAGAGACAGTTATCACATCTGTGGGAGACGAAGAAGGCAGAAGGAGCC
ACGAGTGCATCATCGACGAGGACTGTGGGCCCAGCATGTACTGCCAGTTTGCAGCTTCCAGTAC
ACCTGCCAGCCATGCCGGGGCCAGAGGATGCTTGCACCCGGGACAGTGAGTGCTGTGGAGACCA
GCTGTGTGTCTGGGGTCACTGCACCAAAATGGCCACCAGGGGCGAGCAATGGGACCATCTGTGACA
ACCAGAGGGACTGCCAGCCGGGGCTGTGCTGTGCCTTCCAGAGAGGCTGTGTTCCCTCTGTGTC
ACACCCCTGCCCCGTGGAGGGCGAGCTTTGCCATGACCCCGCCAGCCGGCTTCTGGACCTCATCA
CTGGGAGCTAGAGCCTGATGGAGCCTTGACCGATGCCCTTGTGCCAGTGGCTCTCTGCCAGC
CCCACAGCCACAGCCTGGTGTATGTGTGCAAGCCGACCTTCGTGGGAGCCGTGACCAAGATGGG
GAGATCCTGCTGCCAGAGAGGTCCCCGATGAGTATGAAGTTGGCAGCTTCATGGAGGAGGTGCG
CCAGGAGCTGGAGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCGCTGGGGGAGCCTGCGGCTG
CCGCCGCTGCACTGCTGGGAGGGGAAGAGATTAGATCTGGACCAGGCTGTGGGTAGATGTGCAA
TAGAAATAGCTAATTTATTTCCCCAGGTGTGTGCTTTAGGCGTGGGCTGACCAGGCTTCTTCCCTA
CATCTTCTTCCAGTAAGTTTCCCTCTGGCTTGACAGCATGAGGTGTTGTGCATTTGTTACAGT
CCCCAGGCTGTTCTCCAGGCTTACAGTCTGGTGCTTGGGAGAGTCAGGCAGGGTTAACTGCA
GGAGCAGTTTGGCACCCTGTCCAGATTATTGGCTGCTTTGCCCTTACAGTTGGCAGACAGCCG
TTTGTCTACATGGCTTTGATAATTGTTTGGGGGAGGAGATGAAACAATGTGGAGTCTCCCTC
TGATTGGTTTTGGGGAAATGTGGAGAAGAGTGCCCTGCTTTGCAAACATCAACCTGGCAAAATG
CAACAAATGAATTTCCACGCAGTTCTTCCATGGGCATAGGTAAGCTGTGCCTTCAGCTGTGTC
AGATGAAATGTTCTGTTCACCTGCATTACATGTGTTTATTATCCAGCAGTGTGTGCTCAGCTCC
TACCTCTGTGCCAGGGCAGCATTTCATATCCAAGATCAATTCCCTCTCTCAGCACAGCCTGGGG
AGGGGGTCATTGTTCTCCTCGTCCATCAGGGATCTCAGAGGCTCAGAGACTGCAAGCTGCTTGCC
CAAGTCACACAGCTAGTGAAGACCAGAGCAGTTTCATCTGGTTGTGACTCTAAGCTCAGTGTCT
CTCCACTACCCACACCAGCCTTGGTGCCACCAAAAGTGCTCCCCAAAAGGAGGAGAATGGGAT
TTTTCTTGAGGCATGCACATCTGGAATTAAGGTCAAATAATCTCACATCCCTCTAAAAGTAAA
CTACTGTTAGGAACAGCAGTGTCTCAGAGTGTGGGGCAGCCGCTCTTCTAATGAAGACAATGAT
ATTGACACTGTCCCTCTTTGGCAGTTGCATTAGTAACTTTGAAAGGTATATGACTGAGCGTAGCA
TACAGGTAACTGCAGAAACAGTACTTAGGTAATTGTAGGGCGAGGATTATAAATGAAATTTGC
AAAATCACTTAGCAGCAACTGAAGACAATTATCAACCACGTGGAGAAAATCAAACCGAGCAGGGC
TGTGTGAAACATGGTTGTAATATGCGACTGCGAACTGAACCTACGCCACTCCACAAATGATG
TTTTCAGGTGTATGGACTGTTGCCACCATGTATTCATCCAGAGTCTTAAAGTTTAAAGTTGCA
CATGATTGTATAAGCATGCTTTCTTTGAGTTTTAAATTATGTATAAACATAAGTTGCATTTAGAA
ATCAAGCATAAATCACTTCAACTGCAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 8

MQRLGATLLCLLLAAAVPTAPAPAPTATSAPVKPGPALSYPQEEATLNEMFREVEELMEDTQHKL
RSAVEEMEAEEAAKASSEVNLANLPPSYHNETNTDTKVGNNTIHVHREIHKITNNQTGQMVFSE
TVITSVGDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRGQRMMLCTRDECCGDQLCVWGHC
TKMATRGSNGTICDNQRDCQPLCCAFQRGLLFPVCTPLPVEGELCHDPASRLDLLITWELEPDG
ALDRCPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEVGSFMEEVRQEELEDLE
RSLTEEMALGEPAAAAAALLGGEEI

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 96-100, 106-110, 121-125, 204-208

Casein kinase II phosphorylation site.

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,
327-331

N-myristoylation site.

amino acids 202-208, 217-223

Amidation site.

amino acids 140-144

FIGURE 9

CGGACGCGTGGGCGGACGCGTGGGGGCTGTGAGAAAGTGCCAATAAATACATCATGCAACCCAC
 GGCCACCTTGTGAACCTCCTCGTGCCAGGGCTGATGTGCGTCTTCCAGGGCTACTCATCCAAAG
 GCCTAATCCAACGTTCTGTCTTCAATCTGCAAATCTATGGGGTCTGGGGCTCTTCTGGACCCTT
 AACTGGGTACTGGCCCTGGGCCAATGCGTCCTCGCTGGAGCCTTTGCCTCCTTCTACTGGGCCTT
 CCACAAGCCCCAGGACATCCCTACCTTCCCCTTAATCTCTGCCTTCATCCGCACACTCCGTTACC
 ACACTGGGTCAATTGGCATTGGAGCCCTCATCCTGACCCTTGTGCAGATAGCCGGGTCACTTG
 GAGTATATTGACCACAAGCTCAGAGGAGTGCGAACCCTGTAGCCCGCTGCATCATGTGCTGTTT
 CAAGTGCTGCCTCTGGTGTCTGAAAAATTTATCAAGTTCCTAAACCGCAATGCATACATCATGA
 TCGGCATCTACGGGAAGAATTTCTGTGTCTCAGCCAAAAATGCGTTCATGCTACTCATGCGAAAC
 ATTGTCAGGGTGGTCTCCTGGACAAAGTCACAGACCTGCTGCTGTTCTTTGGGAAGCTGCTGGT
 GGTGGAGGCGTGGGGTCTGTCTTCTTTTTTTTCTCCGGTGCATCCCGGGGCTGGGTAAAG
 ACTTTAAGAGCCCCACCTCAACTATTACTGGCTGCCCATCATGACCTCCATCCTGGGGGCCTAT
 GTCATCGCCAGCGGCTTCTTCAGCGTTTTTCGGCATGTGTGTGGACACGCTCTTCCTCTGCTTCCT
 GGAAGACCTGGAGCGGAACAACGGCTCCCTGGACCGGCCCTACTACATGTCCAAGAGCCTTCTAA
 AGATTCTGGGCAAGAAGAACGAGGCGCCCCCGGACAACAAGAAGAGGAAGAAGTGAACAGCTCCGG
 CCCTGATCCAGGACTGCACCCACCCCCACCGTCCAGCCATCCAACCTCACTTCGCCTTACAGGT
 CTCCATTTTGTGGTAAAAAAGGTTTTAGGCCAGGCGCCGTGGCTCACGCCTGTAATCCAACACT
 TTGAGAGGCTGAGGCGGGCGGATCACCTGAGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTG
 AAACCTCCGTCTCTATTAAAAATACAAAAATTAGCCGAGAGTGGTGGCATGCACCTGTCATCCCA
 GCTACTCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCGGGAGGCAGAGGTTGCAGTGAGCCGA
 GATCGCGCCACTGCACTCCAACCTGGGTGACAGACTCTGTCTCCAAAACAAAACAAAACAAA
 AAGATTTTATTAAAGATATTTTGTAACTC

FIGURE 10

RTRGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVLGFWTL
 NWVLALGQCVLAGAFASFYWAFHKPQDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQIARVIL
 EYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGKNFCVSAKNAEMLLMRN
 IVRVVVLDKVTDLLLFFGKLLVVGGVGVLSFFFFSGRIPGLGKDFKSPHLNYYWLPIMTSILGAY
 VIASGFFSVFGMCVDTLFLCFLEDLERNNGSLDRPYMSKSLKILGKKNEAPPDNKKRKK

Important features:

Transmembrane domains:

amino acids 57-80 (type II), 110-126, 215-231, 254-274

N-glycosylation sites.

amino acids 16-20, 27-31, 289-293

Hypothetical YBR002c family proteins.

amino acids 276-288

Ammonium transporters proteins.

amino acids 204-231

N-myristoylation sites.

amino acids 60-66, 78-84

Amidation site.

amino acids 306-310

FIGURE 11

GCCCCGCGCCCGGCGCCGGGCGCCCGAAGCCGGGAGCCACCGCCATGGGGGCCTGCCTGGGAGCCTGC
TCCCTGCTCAGCTGCGCGTCCTGCCTCTGCGGCTCTGCCCCCTGCATCCTGTGCAGCTGCTGCCCCG
CAGCCGCAACTCCACCGTGAGCCGCTCATCTTCACGTTCTTCCCTCTCCTGGGGGTGCTGGTGTCCA
TCATTATGCTGAGCCCGGGCGTGGAGAGTCAGCTCTACAAGCTGCCCTGGGTGTGTGAGGAGGGGGCC
GGGATCCCCACCGTCCTGCAGGGCCACATCGACTGTGGCTCCCTGCTTGGCTACCGCGCTGTCTACCG
CATGTGCTTCCGCCACGGCGGCCTTCTTCTTCTTCTTTTACCCCTGCTCATGCTCTGCGTGAGCAGCA
GCCGGGACCCCGGGCTGCCATCCAGAATGGGTTTTGGTTCTTTAAGTTCCCTGATCCTGGTGGGCCTC
ACCGTGGGTGCCTTCTACATCCCTGACGGCTCCTTCACCAACATCTGGTTCTACTTCGGCGCTGTGGG
CTCCTTCTTCTCATCTCATCCAGCTGGTGTGCTCATCGACTTTGCGCACTCCTGGAACCAGCGGT
GGCTGGGCAAGGCCGAGGAGTGCATTCCCGTGCCTGGTACGCAGGCCTCTTCTTCTTCACTCTCCTC
TTCTACTTGCTGTGATCGCGGCGTGGCGCTGATGTTTCATGTACTACACTGAGCCAGCGGCTGCCA
CGAGGGCAAGGTCTTCATCAGCCTCAACCTCACCTTCTGTGTCTGCGTGTCCATCGCTGTCTGTCTGC
CCAAGGTCCAGGACGCCAGCCCAACTCGGGTCTGCTGCAGGCCTCGGTATCACCTCTACACCATG
TTTGTACCTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCAACCCCATTTGCCAACCAGCT
GGGCAACGAGACAGTTGTGGCAGGCCCCGAGGGCTATGAGACCCAGTGGTGGGATGCCCGGAGCATTG
TGGGCCTCATCATCTTCTCCTGTGCACCCCTCTTCATCAGTCTGCGCTCCTCAGACCACCGGAGGTG
AACAGCCTGATGCAGACCGAGGAGTGGCCACCTATGCTAGACGCCACACAGCAGCAGCAGCAGGTT
GGCAGCCTGTGAGGGCCGGGCCTTTGACAACGAGCAGGACGGCGTCACCTACAGCTACTCCTTCTTCC
ACTTCTGCCTGGTGTGCGCTCACTGCACGTGATGATGACGCTCACCAACTGGTACAAGCCCGGTGAG
ACCCGAAGATGATCAGCACGTGGACCGCCGTGTGGGTGAAGATCTGTGCCAGCTGGGCAGGGCTGCT
CCTCTACCTGTGGACCTGGTAGCCCCACTCCTCCTGCGCAACCGCGACTTCAGCTTGAGGCAGCCTCA
CAGCCTGCCATCTGGTGCTCCTGCCACCTGGTGCCTCTCGGCTCGGTGACAGCCAACCTGCCCCCTC
CCCACACCAATCAGCCAGGCTGAGCCCCACCCCTGCCCCAGCTCCAGGACCTGCCCTGAGCCGGGC
CTTCTAGTCGTAGTGCTTCAGGGTCCGAGGAGCATCAGGCTCCTGCAGAGCCCCATCCCCCGCCAC
ACCCACACGGTGGAGCTGCCTCTTCCCTTCCCTCCTCCTGTTGCCATACTCAGCATCTCGGATGAA
AGGGCTCCCTTGTCTCAGGCTCCACGGGAGCGGGGCTGCTGGAGAGAGCGGGGAACCTCCACCACAG
TGGGGCATCCGGCACTGAAGCCCTGGTGTCTGCTCAGTCCCCCAGGGGACCCTGCCCCCTTCTCTG
GACTTCGTGCCTTACTGAGTCTCTAAGACTTTTTCTAATAAACAAGCCAGTGCCTGTAAAAAAA

$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}} \right) = \frac{\partial L}{\partial x}$

Signal sequence:

Transmembrane domains:

Page 162 of 318

FIGURE 13

CGGGCCAGCCTGGGGCGGCCGCCAGGAACCAACCCGTTAAGGTGTCTTCTCTTTAGGGATGGTGA
 GGTGGAAGAAAGACTCCTGTAACCCCTCCTCCAGGATGAACCAACCTGCCAGAAGACATGGAGAACG
 CTCTCACCGGGAGCCAGAGCTCCCATGCTTCTCTGCGCAATATCCATTCCATCAACCCACACAA
 CTCATGGCCAGGATTGAGTCCATGAAGGAAGGAAAAGAAAGGCATATCTGATGTCAGGAGGAC
 TTTCTGTTTGTGTCACCTTTGACCTCTTATTCGTAACATTACTGTGGATAATAGAGTTAAATG
 TGAATGGAGGCATTGAGAACACATTAGAGAAGGAGGTGATGCAGTATGACTACTATTCTTCATAT
 TTTGATATATTTCTTCTGGCAGTTTTTCGATTAAAGTGTAATACTTGCAATATGCTGTGTGCAG
 ACTGCGCCATTGGTGGGCAATAGCGTTGACAACGGCAGTGACCAGTGCCTTTTTACTAGCAAAAG
 TGATCCTTTTCAAGCTTTTCTCTCAAGGGGCTTTTGGCTATGTGCTGCCCATCATTTTCATTATC
 CTTGCTCGGATTGAGACGTGGTTCCTGGATTTCAAAGTGTTACCTCAAGAAGCAGAAGAAGAAAA
 CAGACTCCTGATAGTTCAGGATGCTTCAGAGAGGGCAGCACTTATACCTGGTGGTCTTTCTGATG
 GTCAGTTTTATTCCCCTCCTGAATCCGAAGCAGGATCTGAAGAAGCTGAAGAAAAACAGGACAGT
 GAGAAACCACTTTTAGAACTATGAGTACTACTTTTGTTAAATGTGAAAAACCCCTCACAGAAAGTC
 ATCGAGGCAAAAAGAGGCAGGCAGTGGAGTCTCCCTGTGACAGTAAAGTTGAAATGGTGACGTC
 CACTGCTGGCTTTATTGAACAGCTAATAAAGATTTATTTATTGTAATACCTCACAAACGTTGTAC
 CATATCCATGCACATTTAGTTGCCTGCCTGTGGCTGGTAAGGTAATGTCATGATTATCCTCTCT
 TCAGTGAGACTGAGCCTGATGTGTTAACAAATAGGTGAAGAAAGTCTTGTGCTGTATTCTTAATC
 AAAAGACTTAATATATTGAAGTAACACTTTTTTAGTAAGCAAGATACCTTTTTATTTCATTCAC
 AGAATGGAATTTTTTTGTTTCATGTCTCAGATTATTTTGTATTTCTTTTTTAACACTCTACATT
 TCCCTTGTTTTTTAACTCATGCACATGTGCTCTTTGTACAGTTTTTAAAAGTGTAATAAAATCTG
 ACATGTCAATGTGGCTAGTTTTATTTTCTTGTGTTTGCATTATGTGTATGGCCTGAAGTGTGGA
 CTTGCAAAAGGGGAAGAAAGGAATTGCGAATACATGTAAATGTCACCAGACATTTGTATTATTT
 TTATCATGAAATCATGTTTTTCTCTGATTGTTCTGAAATGTTCTAAATACTCTTATTTTGAATGC
 ACAAAATGACTTAAACCATTATATCATGTTTCCTTTGCGTTTCAGCCAATTCAATTAAATGAA
 CTAATTAATAA

FIGURE 14

MNHLPEDMENALTGSQSSHASLRNIHSINPTQLMARIESYEGREKKGISDVRRTFCLEVTFDLLF
VTLLWIIELNVNGGIENTLEKEVMQYDYYSSYFDIFLLAVFRFKVLILAYAVCRLRHWWAIALTT
AVTSAFLLAKVILSKLFSQGAFGYVLPPIISFILAWIETWFLDFKVLQPQAEENRLLIVQDASER
AALIPGGLSDGQFYSPPESEAGSEEAEKQDSEKPLLEL

Important features of the protein:

Signal peptide:

amino acids 1-20

Transmembrane domains:

amino acids 54-72, 100-118, 130-144, 146-166

N-myristoylation sites.

amino acids 14-20, 78-84, 79-85, 202-208, 217-223

FIGURE 15

ACTCGAACGCAGTTGCTTCGGGACCCAGGACCCCTCGGGCCCGACCCGCCAGGAAAGACTGAGG
 CCGCGGCCTGCCCGCCCGGCTCCCTGCGCCGCCGCCCTCCCGGACAGAAGATGTCCTCCAG
 GGTCCCTCTGCTGCTGCCGCTGCTCCTGCTACTGGCCCTGGGGCTGGGGTGACAGGGCTGCCCAT
 CCGGCTGCCAGTGACGCCAGCCACAGACAGTCTTCTGCACTGCCCGCCAGGGGACCACGGTGCCC
 CGAGACGTGCCACCCGACACGGTGGGGCTGTACGTCTTTGAGAACGGCATCACCATGCTCGACGC
 AGGCAGCTTTGCCGGCTGCCGGGCTGCAGCTCCTGGACCTGTACAGAACAGATCGCCAGCC
 TGCCAGCGGGGTCTTCCAGCCACTCGCCAACCTCAGCAACCTGGACCTGACGGCCAACAGGCTG
 CATGAAATCACCATGAGACCTTCCGTGGCCTGCGGCGCCTCGAGCGCCTCTACCTGGGCAAGAA
 CCGCATCCGCCACATCCAGCCTGGTGCTTCGACACGCTCGACCGCCTCCTGGAGCTCAAGCTGC
 AGGACAACGAGCTGCGGGCACTGCCCGCTGCGCCTGCCCGCCTGCTGCTGCTGGACCTCAGC
 CACAACAGCCTCCTGGCCTGGAGCCCGGCATCCTGGACACTGCCAACGTGGAGGCGCTGCGGCT
 GGCTGGTCTGGGGCTGCAGCAGCTGGACGAGGGGCTCTTCAGCCGCTTGCGCAACCTCCACGACC
 TGGATGTGTCCGACAACCAGCTGGAGCGAGTGCCACCTGTGATCCGAGGCTCCGGGGCTGACG
 CGCCTGCGGCTGGCCGGCAACACCGCATTTGCCAGCTGCGGCCGAGGACCTGGCCGGCTGGC
 TGCCCTGCAGGAGCTGGATGTGAGCAACCTAAGCCTGCAGGCCCTGCCTGGCGACCTCTCGGGCC
 TCTTCCCCCGCCTGCGGCTGCTGGCAGCTGCCCGCAACCCCTTCAACTGCGTGTGCCCTTGAGC
 TGGTTTGGCCCCCTGGGTGCGCGAGAGCCACGTACACTGGCCAGCCCTGAGGAGACGCGCTGCCA
 CTTCCCGCCCAAGAACGCTGGCCGGCTGCTCCTGGAGCTTGACTACGCCGACTTTGGCTGCCAG
 CCACCACCACACAGCCACAGTGCCCAACACGAGGCCGCTGGTGCGGGAGCCACAGCCTTGTCT
 TCTAGCTTGGCTCCTACCTGGCTTAGCCCCACAGCGCCGGCCACTGAGGCCCCAGCCCCCCTC
 CACTGCCCCACCGACTGTAGGGCTGTCCCCAGCCCCAGGACTGCCACCGTCCACCTGCCTCA
 ATGGGGGCACATGCCACCTGGGGACACGGCACCACTGGCGTGCTTGTGCCCGAAGGCTTACG
 GGCTGTACTGTGAGAGCCAGATGGGGCAGGGGACACGGCCAGCCCTACACCAGTCACGCCAG
 GCCACCACGGTCCCTGACCCTGGGCATCGAGCCGGTGAGCCCCACCTCCCTGCGCGTGGGGCTGC
 AGCGCTACCTCCAGGGGAGCTCCGTGCAGCTCAGGAGCCTCCGTCTCACCTATCGCAACCTATCG
 GGCCCTGATAAGCGGCTGGTGACGCTGCGACTGCCTGCCTCGCTCGCTGAGTACACGGTCACCCA
 GCTGCGGCCCCAACGCCACTTACTCCGTCTGTGTCATGCCTTTGGGGCCCGGGGGTGGCGGAGG
 GCGAGGAGGCTGCGGGGAGGGCCATACACCCAGCCGTCCACTCCAACCACGCCCCAGTCACC
 CAGGCCCCGCGAGGGCAACCTGCCGCTCCTCATTGCGCCCGCCCTGGCCGCGGTGCTCCTGGCCGC
 GCTGGCTGCGGTGGGGGAGCCTACTGTGTGCGGCGGGGGCGGGCCATGGCAGCAGCGGCTCAGG
 ACAAGGGGCAGGTGGGGCCAGGGGCTGGGCCCTGGAACCTGGAGGGAGTGAAGGTCCCCTTGGAG
 CCAGGCCCGAAGGCAACAGAGGGCGGTGGAGAGGCCCTGCCAGCGGGTCTGAGTGTGAGGTGCC
 ACTCATGGGCTTCCCAGGGCTTGGCCTCCAGTCACCCCTCCACGCAAGCCCTACATCTAAGCCA
 GAGAGAGACAGGGCAGCTGGGGCCGGGCTCTCAGCCAGTGAGATGGCCAGCCCCCTCCTGCTGCC
 ACACCACGTAAGTTCTCAGTCCCAACCTCGGGGATGTGTGCAGACAGGGCTGTGTGACCACAGCT
 GGGCCTGTTCCTCTGGACCTCGGTCTCCTCATCTGTGAGATGCTGTGGCCAGCTGACGAGCC
 CTAACGTCCCCAGAACCAGTGCCTATGAGGACAGTGTCCGCCCTGCCCTCCGCAACGTGCAGTC
 CCTGGGCACGGCGGGCCCTGCCATGTGCTGGTAACGCATGCCTGGGTCTGCTGGGCTCTCCAC
 TCCAGCGCGACCCCTGGGGGCCAGTGAAGGAAGCTCCCGGAAAGAGCAGAGGGAGAGCGGCTAGGC
 GGCTGTGTGACTCTAGTCTTGGCCCCAGGAAGCGAAGGAACAAAAGAACTGGAAGGAAGATGC
 TTTAGGAACATGTTTGTCTTTTAAATATATATATTTATAAGAGATCCTTTCCCATTTATTCT
 GGAAGATGTTTTTCAACTCAGAGACAAGGACTTTGGTTTTTGTAAAGACAACGATGATATGAA
 GGCTTTTGTAAAGAAAAATAAAGATGAAGTGTGAAA

FIGURE 16

MCSRVP LLLPL LLLLLALGPGVQGCPSGCQCSQPQT V FCTARQGT TVPRDVPPDTVGLYVFENGIT
 MLDAGS FAGLPGLQLLDLSQNQIASLPSGVFQPLANLSNLDLTANRLHEITNETFRGLRRLERLY
 LGKNRIRHIQPGAFDTLDRLELKLQDNELRALPPLRLPRLLLLDLSHNSLLALEPGILDTANVE
 ALRLAGLGLQQLDEGLFSRLRNLDLDVSDNQLERVPPVIRGLRGLTRLRLAGNTRIAQLRPEDL
 AGLAALQELDVSNLSLQALPGDLSGLFPRLRLAAARNPFNCVCPLSWFGPWVRESHVTLASPEE
 TRCHFPPKNAGRLLLELDYADFGCPATTTTATVPTTRPVVREPTALSSSLAPTWLSPTAPATEAP
 SPPSTAPPTVGPVPQPQDCPPSTCLNGGTCHLGRHHLACLCEGFTGLYCESQMGQGTRPSPTP
 VTPRPPRS LTLGIEPVSPTS LRVGLQRYLQGSSVQLRSLRLTYRNLSGPDKRLVTLRLPASLA EY
 TVTQLRP NATYSVCVMPLGPGRVPEGEEACGEAHTPPAVHSNHAPVTQAREGNLPLLIAPALAAV
 LLAALAAVGAAYCVRGRAMAAAAQDKGQVGPAGPLELEGVKVPLEPGPKATEGGGEALPSGSE
 CEVPLMGFPGLQSP L HAKPYI

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 579-599

EGF-like domain cysteine pattern signature.

amino acids 430-442

Leucine zipper pattern.

amino acids 197-219, 269-291

N-glycosylation sites.

amino acids 101-105, 117-121, 273-277, 500-504, 528-532

Tyrosine kinase phosphorylation sites.

amino acids 124-131, 337-345

N-myristoylation sites.

amino acids 23-29, 27-33, 70-76, 142-148, 187-193, 348-354,
 594-600, 640-646

FIGURE 17

GCAGCGGCGAGGCGGCGGTGGTGGCTGAGTCCGTGGTGGCAGAGGCGAAGGCGACAGCTCATGCG
 GGTCCGGATAGGGCTGACGCTGCTGCTGTGTGCGGTGCTGCTGAGCTTGGCCTCGGCGTCCTCGG
 ATGAAGAAGGCAGCCAGGATGAATCCTTAGATTCCAAGACTACTTTGACATCAGATGAGTCAGTA
 AAGGACCATACTACTGCAGGCAGAGTAGTTGCTGGTCAAATATTTCTTGATTGAGAAGAATCTGA
 ATTAGAATCCTCTATTCAAGAAGAGGAAGACAGCCTCAAGAGCCAAGAGGGGAAAGTGTACAG
 AAGATATCAGCTTTCTAGAGTCTCCAAATCCAGAAAACAAGGACTATGAAGAGCCAAAGAAAGTA
 CGGAAACCAGCTTTGACCGCCATTGAAGGCACAGCACATGGGGAGCCCTGCCACTTCCCTTTTCT
 TTTCTAGATAAGGAGTATGATGAATGTACATCAGATGGGAGGGAAGATGGCAGACTGTGGTGTG
 CTACAACCTATGACTACAAAGCAGATGAAAAGTGGGGCTTTTGTGAAACTGAAGAAGAGGCTGCT
 AAGAGACGGCAGATGCAGGAAGCAGAAATGATGTATCAAACCTGGAATGAAAATCCTTAATGGAAG
 CAATAAGAAAAGCCAAAAAAGAGAAGCATATCGGTATCTCCAAAAGGCAGCAAGCATGAACCATA
 CCAAAGCCCTGGAGAGAGTGTATATGCTCTTTTATTTGGTGATTACTTGCCACAGAATATCCAG
 GCAGCGAGAGAGATGTTGAGAAGCTGACTGAGGAAGGCTCTCCAAGGGACAGACTGCTCTTGG
 CTTTCTGTATGCCTCTGGACTTGGTGTAAATCAAGTCAGGCAAAGGCTCTTGTATATTATACAT
 TTGGAGCTCTTGGGGCAATCTAATAGCCCACATGGTTTTGGTAAGTAGACTTTAGTGGAAGGCT
 AATAATATTAACATCAGAAGAATTTGTGGTTTATAGCGGCCACAACCTTTTTCAGCTTTCATGATC
 CAGATTTGCTTGTATTAAGACCAATATTCAGTTGAACTTCCTTCAAATCTTGTAAATGGATAT
 AACACATGGAATCTACATGTAAATGAAAGTTGGTGGAGTCCACAATTTTCTTTAAATGATTAG
 TTTGGCTGATTGCCCCATAAAAGAGAGATCTGATAAATGGCTCTTTTAAATTTTCTCTGAGTTG
 GAATTGTCAGAATCATTTTTTACATTAGATTATCATAATTTTAAAAATTTTCTTTAGTTTTCA
 AAATTTTGTAATGGTGGCTATAGAAAAACAACATGAAATATTATACAATATTTTGCAACAATGC
 CCTAAGAATTGTTAAATTCATGGAGTTATTTGTGCAGAATGACTCCAGAGAGCTCTACTTTCTG
 TTTTTTACTTTTCATGATTGGCTGTCTTCCATTTATTCTGGTCATTTATTGCTAGTGACACTGT
 GCCTGCTTCCAGTAGTCTCATTTTCCCTATTTTGCTAATTTGTTACTTTTTCTTGCTAATTTGG
 AAGATTAACCTCATTTTTTAATAAATATATGTCTAAGATTAAAAAAAAAAAAAAAAAAAAAAAAA
 AA

FIGURE 21

CTGAGGCGGCGGTAGCATGGAGGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTTGTGCTCGGCG
 CACTCGCTTTCCAGCACCTCAACACGGACTCGGACACGGAAGGTTTTCTTCTTGGGGAAGTAAAA
 GGTGAAGCCAAGAACAGCATTACTGATTCCCAAATGGATGATGTTGAAGTTGTTTATACAATTGA
 CATTGAGAAATATATTCCATGCTATCAGCTTTTCTAGCTTTTATAATTCTTCAGGCGAAGTAAATG
 AGCAAGCACTGAAGAAAATATTATCAAATGTCAAAAAGAAATGTGGTAGGTTGGTACAAATTCCGT
 CGTCATTGAGATCAGATCATGACGTTTAGAGAGAGGCTGCTTCACAAAACCTTGCAGGAGCATT
 TTCAAACCAAGACCTTGTTTTCTGCTATTACACCAAGTATAATAACAGAAAGCTGCTCTACTC
 ATCGACTGGAACATTCCCTTATATAAACCTCAAAAAGGACTTTTTCACAGGGTACCTTTAGTGGTT
 GCCAATCTGGGCATGCTGAACAACCTGGGTATATAAACTGTATCAGGTTCCCTGTATGTCCACTGG
 TTTTAGCCGAGCAGTACAAACACACAGCTCTAAATTTTTGAAGAAGATGGATCCTTAAAGGAGG
 TACATAAGATAAATGAAATGTATGCTTCATTACAAGAGGAATTAAGAGTATATGCAAAAAGTG
 GAAGACAGTGAACAAGCAGTAGATAAACTAGTAAAGGATGTAAACAGATTAAAACGAGAAATTGA
 GAAAAGGAGAGGAGCACAGATTGAGGCAGCAAGAGAGAAGAACATCCAAAAGACCCTCAGGAGA
 ACATTTTTCTTTGTCAGGCATTACGGACCTTTTTTCCAAATCTGAATTTCTTCATTGATGTT
 ATGTCTTTAAAAATAGACATGTTTCTAAAAGTAGCTGTAACACCAACCACCATCTCGATGTAGT
 AGACAACTCTGACCTTAATGGTAGAACACACTGACATTCCCTGAAGCTAGTCCAGCTAGTACACCAC
 AAATCATTAAGCATAAAGCCTTAGACTTAGATGACAGATGGCAATTCAGAGATCTCGGTTGTTA
 GATACACAAGACAAACGATCTAAAGCAAATACTGGTAGTAGTAACCAAGATAAAGCATCCAAAAT
 GAGCAGCCCGAGAAACAGATGAAGAAATTGAAAAGATGAAGGGTTTTGGTGAATATTCACGGTCTC
 CTACATTTTGATCCTTTTAACCTTACAAGGAGATTTTTTATTTGGCTGATGGGTAAAGCCAAAC
 ATTTCTATTGTTTTTACTATGTTGAGCTACTTGACAGTAAGTTCATTTGTTTTTACTATGTTTACC
 TGTTTGCAGTAATACACAGATAACTCTTAGTGCAATTTACTTCACAAAGTACTTTTTCAAACATCA
 GATGCTTTTATTTCCAAACCTTTTTTTCACCTTCACTAAGTTGTTGAGGGGAAGGCTTACACAG
 ACACATTCTTTAGAATTGGAAGAGTGAGACCAGGCACAGTGGCTCACACCTGTAATCCCAGCACT
 TAGGGAAGACAAGTCAGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCCTGGGCAACGTATT
 GAGACCATGTCTATTAAAAAATAAAATGAAAAGCAAGAATAGCCTTATTTCAAATATGGAAA
 GAAATTTATATGAAAATTTATCTGAGTCATTAAATTTCTCCTAAGTGATACTTTTTTAGAAGTA
 CATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCAATAAATTTGCAAAACATCATCT
 AAAATTTAAAAA

FIGURE 22

MEGESTSAVLSGFVLGALAFQHLNTSDTEGFLGGEVKGEAKNSITDSQMDDVEVVYTIDIQKYI
 PCYQLFSFYNSSGEVNEQALKKILSNVKKNVVGWYKFRRHSDQIMTFRERLLHKNLQEHFSNQDL
 VFLLLTPSIITESCSTHRLEHSLYKPQKGLFHRVPLVVANLGMSEQLGYKTVSGSCMSTGFSRAV
 QTHSSKFFEEGSLKEVHKINEMYASLQEEELKSICKKVEDSEQAVDKLVKDVNRLKREIEKRRGA
 QIQAAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVSKSSCYNHHLDVVDNLTL
 MVEHTDIPEASPASTPQIIKHKALDLDDRWFKRSRLDQDKRSKANTGSSNQDKASKMSSPET
 DEEIEKMKGFGEYSRSPTF

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation sites.

amino acids 75-79, 322-326

N-myristoylation site.

amino acids 184-154

Growth factor and cytokines receptors family.

amino acids 134-150

FIGURE 23

GGCACAGCCGCGCGGCGGAGGGCAGAGTCAGCCGAGCCGAGTCCAGCCGACGAGCGGACAGCGCAGGGCAGCCCAA
GCAGCGCGCAGCGAACGCCCCGCGCCGCCACACCCTCTGCGGTCCCCGCGGCGCTGCCACCCTTCCCTCCTTCCCC
GCGTCCCCGCTCGCCGCGCAGTCAGCTTGCCGGGTTGCTGCCCGCGGAAACCCGAGGTACCAGCCGCGCCTCT
GCTTCCCTGGGCGCGCGCCGCTCCACGCCCTCCTTCTCCCTGGCCGCGCCTGGCACCGGGGACCGTTGCCTGA
CGCGAGGCCAGCTCTACTTTTCGCCCGCGGTCTCCTCCGCTGCTCGCTCTTCCACCAACTCCAACCTCTCTCCC
TCCAGTCCACTCGTAGTCCCCGACTCCGCCAGCCCTCGGCCCGCTGCCGTAGCGCGCTTCCCGTCCGGTCCCAAA
GGTGGGAACGCGTCCGCCCGGCGCCGACCATGCGACGGTTCCGGCTTGCCCGCGCTTCTCTGCACCCTGGCAGTGCTC
AGCGCGCGCTGCTGGCTGCCGAGCTCAAGTCGAAAAGTTGCTCGGAAGTGCGACGCTTTACGTGTCCAAAGGCTTC
AACAGAAGCATGCCCCCTCCACGAGATCAACGGTGATCATTTGAAGATCTGTCCCAGGGTTCTACCTGCTGCTCT
CAAGAGATGGAGGAGAAGTACAGCCTGCAAAGTAAAGATGATTTCAAAGTGTGGTCAGCGAACAGTGAATCATTTG
CAAGCTGTCTTGTCTTACGTTACAAGAAGTTGATGAATCTTCAAAGAAGTACTTGAAAATGCAGAGAAATCCCTG
AATGATATGTTTGAAGACATATGGCCATTTATACATGCAAATCTGAGCTATTTAAAGATCTCTTCGTAGAGTTG
AAACGTTACTACGTGGTGGGAAATGTGAACCTGGAAGAAATGCTAAATGACTTCTGGGCTCGCCTCCTGGAGCGGATG
TTCCGCTGCTGAACCTCCAGTACCCTTTACAGATGAGTATCTGGAATGTGTGAGCAAGTATACGGAGCAGCTGAAG
CCCTTCGGAGATGTCCTCGCAAATGAAGCTCCAGGTTACTCGTGTCTTTGTAGCAGCCCGTACTTTCGCTCAAGGC
TTAGCGGTTGCGGGAGATGCTGTGAGCAAGGTCTCCGTGGTAAACCCACAGCCCGAGTGTACCCATGCCCTGTTGAAG
ATGATCTACTGCTCCCACTGCCGGGTCTCGTGACTGTGAAGCCATGTTACAAGTACTGCTCAAACATCATGAGAGGC
TGTTTGGCAACCAAGGGGATCTCGATTTTGAATGGAACAATTTATAGATGCTATGCTGATGTTGGCAGAGAGGCTA
GAGGGTCTTTCAACATTGAATCGGTGATGGATCCCATCGATGTGAAGATTTCTGATGCTATTATGAACATGCAAGAT
AATAGTGTCTCAAGTGTCTCAGAAGGTTTTCCAGGGATGTGGACCCCCAAGCCCTCCAGCTGGACGAATTTCTCGT
TCCATCTCTGAAAGTGCCCTTCAGTGCTCGCTTCAGACCACATCACCCGAGGAACGCCAACCCACAGCAGCTGGCACT
AGTTTGGACCGACTGGTTACTGATGTCAAGGAGAACTGAAACAGGCCAAGAAATTTCTGGTCTCCCTTCCGAGCAAC
GTTTGAACGATGAGAGGATGGCTGCAGGAAACGGCAATGAGGATGACTGTTGGAATGGGAAAGGCAAAAGCAGGTAC
CTGTTTGCAGTGACAGGAAATGGATTAGCCAACAGGGCAACAACCCAGAGGTCCAGGTTGACACCAGCAAACAGAC
ATACTGATCCTTCGTCAAATCATGGCTCTTCGAGTGATGACCAGCAAGATGAAGAATGCATACAATGGGAACGACGTG
GACTTCTTTGATATCAGTGATGAAAGTAGTGGAAGGAAGTGAAGTGGCTGTGAGTATCAGCAGTGCCCTTCAGAG
TTTGACTACAATGCCACTGACCATGCTGGGAAGAGTGCCAATGAGAAAGCCGACAGTGTGGTGTCCGCTCCTGGGGCA
CAGGCTACCTCCTCACTGTCTTCTGCATCTTGTCTGGTTATGCAGAGAGAGTGGAGATTAATCTCAAACCTGTGAG
AAAAAGTGTTTCATCAAAAAGTTAAAAGGCACCAAGTTATCACTTTTCTACCATCCTAGTGACTTTGCTTTTAAATGAA
TGGACAACAATGTACAGTTTTTACTATGTGGCCACTGGTTTAAAGAGTGTGACTTTGTTTTCTATTCAAGTTTTGGG
AGGAAAAGGGACTGTGCATTGAGTTGGTTCCCTGCTCCCCAAACCATGTTAAACGTGGCTAACAGTGTAGGTACAGAA
CTATAGTTAGTTGTGCATTTGTGATTTTATCACTCTATTTTGTGTTGATGTTTTTTCTCATTTTCGTTTGTGGGTT
TTTTTTTCCAACTGTGATCTCGCCTTGTTCCTTACAAGCAAACAGGGTCCCTTCTTGGCAGTAACATGTACGTATT
CTGAAATATTAATAGCTGTACAGAAGCAGGTTTATTTATCATGTTATCTTATTAAGAAAAAGCCAAAAAGC

FIGURE 24

MARFGLPALLCTLAVLSAALLAAELKSKSCSEVRRLYVSKGFNKNDAPLHEINGDHLKICPQGST
CCSQEMEKEYSLQSKDDFKSVVSEQC�HNLQAVFASRYKKFDEFFKELLENAEKSLNDMFVKTYGH
LYMQNSELFKDLFVELKRYVVGNVNLEEMLNDFWARLLERMFRLVNSQYHFTDEYLECVSKYTE
QLKPFQDVPRKLLQVTRAFVAARTFAQGLAVAGDVVSKVSVVNPTAQCTHALLKMIYCSHCRL
VTVKPCYNYCSNIMRGCLANQGDLDFFEWNNFIDAMLMVAERLEGPFNIESVMDPIDVKISDAIMN
MQDNSVQVSQKVFQCGGPPKPLPAGRISRSISESAFSARFRPHHPEERPTTAAGTSLDRLVTDVK
EKLKQAKKFWSSLPSNVCNDERMAAGNGNEDDCWNGKGKSRYLFAVTGNGLANQGNNPEVQVDT
KPDILILRQIMALRVMTSKMKNAYNGNDVDFDISDESSGEGSGSGCEYQQCPSEFDYNATDHAG
KSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

Important features:

Signal peptide:

amino acids 1-22

ATP/GTP-binding site motif A (P-loop).

amino acids 515-524

N-glycosylation site.

amino acids 514-518

Glycosaminoglycan attachment sites.

amino acids 494-498, 498-502

N-myristoylation sites.

amino acids 63-69, 224-230, 276-282, 438-444, 497-503, 531-537

Glypicans proteins.

amino acids 54-75, 105-157, 238-280, 309-346, 423-460, 468-506

FIGURE 25

CTCGCCCTCAAATGGGAACGCTGGCCTGGGACTAAAGCATAGACCACCAGGCTGAGTATCCTGAC
 CTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTTCCATTATATTCTTCAAGCAACT
 TACAGCTGCACCGACAGTTGCGATGAAAGTTCTAATCTCTTCCCTCCTCCTGTTGCTGCCACTAA
 TGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGGTGCCCAGAGGCCACAGGGACCGAGGC
 CAGGCTTCTAGGAGATGGCTCCAGGAAGGCGGCCAAGAATGTGAGTGCAAAGATTGGTTCTGAG
 AGCCCCGAGAAGAAAATTCATGACAGTGTCTGGGCTGCCAAGAAGCAGTGCCCTGTGATCATT
 TCAAGGGCAATGTGAAGAAAACAAGACACCAAAGGCACCACAGAAAGCCAAACAAGCATTCCAGA
 GCCTGCCAGCAATTTCTCAAACAATGTCAGCTAAGAAGCTTTGCTCTGCCTTTGTAGGAGCTCTG
 AGCGCCCACTCTTCCAATTAAACATTCTCAGCCAAGAAGACAGTGAGCACACCTACCAGACACTC
 TTCTTCTCCCACCTCACTCTCCCACTGTACCCACCCCTAAATCATTCCAGTGCTCTCAAAAAGCA
 TGTTTTTCAAGATCATTTTGTGTTGCTCTCTCTAGTGTCTTCTTCTCTCGTCAGTCTTAGCCT
 GTGCCCTCCCCTTACCCAGGCTTAGGCTTAATTACCTGAAAGATTCCAGGAACTGTAGCTTCCT
 AGCTAGTGTCAATTAACCTTAAATGCAATCAGGAAAGTAGCAAACAGAAGTCAATAAATATTTTT
 AAATGTCAAAAAAAAAAAAAAAAAA

FIGURE 26

MKVLISLLLLLPLMLSMVSSSLNPGVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRRKFM
TVSGLPKKQPCDHFKGNVKKTRHQRHHRKPNKHSRACQQFLKQCQLRSFALPL

Important features:

Signal peptide:

amino acids 1-22

N-myristoylation sites.

amino acids 27-33, 46-52

FIGURE 27

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTGCCCCAGCGGAAGCACAGCTCAG
 AGCTGGTCTGCCATGGACATCCTGGTCCCCTCCTGCAGCTGCTGGTGCTGCTTCTTACCCTGCC
 CCTGCACCTCATGGCTCTGCTGGGCTGCTGGCAGCCCCCTGTGCAAAAGCTACTTCCCCTACCTGA
 TGGCGTGCTGACTCCCAAGAGCAACCGCAAGATGGAGAGCAAGAAACGGGAGCTCTTCAGCCAG
 ATAAAGGGGCTTACAGGAGCCTCCGGGAAAGTGGCCCTACTGGAGCTGGGCTGCGGAACCGGAGC
 CAACTTTCAGTTCTACCCACCGGGCTGCAGGGTCACCTGCCTAGACCCAAATCCCCACTTTGAGA
 AGTTCTTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATATGAGCGGTTTGTGGTGGCTCCT
 GGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGTGGTGGTCTGCACTCTGGTGCTGTG
 CTCTGTGCAGAGCCCAAGGAAGGTCCTGCAGGAGGTCCGGAGAGTACTGAGACCGGGAGGTGTGC
 TCTTTTCTGGGAGCATGTGGCAGAACCATATGGAAGCTGGGCCTTCATGTGGCAGCAAGTTTTTC
 GAGCCACCTGGAAACACATTGGGGATGGCTGCTGCCTCACCAGAGAGACCTGGAAGGATCTTGA
 GAACGCCAGTTCTCCGAAATCCAAATGGAACGACAGCCCCCTCCCTTGAAGTGGCTACCTGTTG
 GGCCCCACATCATGGGAAAGGCTGTCAAACAATCTTTCCCAAGCTCCAAGGCACCTCATTTGCTCC
 TTCCCCAGCCTCCAATTAGAACAAGCCACCCACCAGCCTATCTATCTTCCACTGAGAGGGACCTTA
GCAGAATGAGAGAAGACATTCTGTACCACCTACTAGTCCCTCTCTCCCCAACCTCTGCCAGGGC
 AATCTCTAACTTCAATCCCGCCTTCGACAGTGAAAAAGCTCTACTTCTACGCTGACCCAGGGAGG
 AAACACTAGGACCCTGTTGTATCCTCAACTGCAAGTTTCTGGACTAGTCTCCCAACGTTTGCCTC
 CCAATGTTGTCCCTTTCCTTCGTTCCCATGGTAAAGCTCCTCTCGCTTTCCTCCTGAGGCTACAC
 CCATGCGTCTCTAGGAAGTGGTCACAAAAGTCATGGTGCCTGCATCCCTGCCAAGCCCCCTGAC
 CCTCTCTCCCCACTACCACCTTCTTCCTGAGCTGGGGGCACCAGGGAGAATCAGAGATGCTGGGG
 ATGCCAGAGCAAGACTCAAAGAGGCAGAGGTTTGTCTCAAATATTTTAAATAAATAGACGAA
 ACCACG

FIGURE 28

MDILVPLLQLLVLLLTLP LHLMA LLGCWQPLCKSYFPYLM AVLTPKSNRKMESKKRELFSQIKGL
TGASGKVALLELGCGTGANFQFYPPGCRVTC LDPNPHFEKFLTKSMAENRHLQYERFVVAPGEDM
RQLADGSMDVVVCTLVLC SVQSPRKVLQEVRRVLRPGGV LFFWEHVAEPYGSWAFMWQQVF EPTW
KHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWLPVGP HIMGKAVKQSFSPSSKALICSFP SL
QLEQATHQPIYLPLRGT

Important features:

Signal peptide:

amino acids 1-23

Leucine zipper pattern.

amino acids 10-32

N-myristoylation sites.

amino acids 64-70, 78-84, 80-86, 91-97, 201-207

FIGURE 30

MLLLTLLLLLLLLLKGSCLWGLVGAQKVSSATDAPIRDWAFPPSFLCLLPHRPAMTCSQAQPRG
EGEKVGDG

Important features:

Signal peptide:

amino acids 1-15

Growth factor and cytokines receptors family:

amino acids 3-18

FIGURE 31

GTTTGAATTCCTTCAACTATACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCAGTT
CCTCCAAGCAAGTCATTTCCCTTATTTAACCGATGTGTCCCTCAAACACCTGAGTGCTACTCCCT
ATTTGCATCTGTTTTGATAAATGATGTTGACACCCCTCCACCGAATTCTAAGTGGAATCATGTCGG
GAAGAGATACAATCCTTGGCCTGTGTATCCTCGCATTAGCCTTGTCTTTGGCCATGATGTTTACC
TTCAGATTCATCACCACCCTTCTGGTTCACATTTTCATTTCATTGGTTATTTTGGGATTGTTGTT
TGTCTGCGGTGTTTTATGGTGGCTGTATTATGACTATACCAACGACCTCAGCATAGAATTGGACA
CAGAAAGGGAAAATATGAAGTGCCTGCTGGGGTTTGCTATCGTATCCACAGGCATCACGGCAGTG
CTGCTCGTCTTGATTTTTGTCTCAGAAAGAGAATAAAATTGACAGTTGAGCTTTTCCAAATCAC
AAATAAAGCCATCAGCAGTGCTCCCTTCCTGCTGTTCCAGCCACTGTGGACATTTGCCATCCTCA
TTTTCTTCTGGGTCTCTGGGTGGCTGTGCTGCTGAGCCTGGGAAGTGCAGGAGCTGCCAGGTT
ATGGAAGGCGGCCAAGTGAATATAAGCCCCTTTCGGGCATTGGGTACATGTGGTCGTACCATTT
AATTGGCCTCATCTGGACTAGTGAATTCATCCTTGCCTGCCAGCAAATGACTATAGCTGGGGCAG
TGGTTACTTGTATTTCAACAGAAGTAAAAATGATCCTCCTGATCATCCCATCCTTTCGTCTCTC
TCCATTCTCTTCTTCTACCATCAAGGAACCGTTGTGAAAGGGTCATTTTAAATCTCTGTGGTGAG
GATTCCGAGAATCATGTGCATGTACATGCAAAACGCACCTGAAAGAACAGCAGCATGGTGCATTGT
CCAGGTACCTGTTCCGATGCTGCTACTGCTGTTTCTGGTGTCTTGACAAATACCTGCTCCATCTC
AACCAGAATGCATATACTACAAGTCTATTAATGGGACAGATTTCTGTACATCAGCAAAAGATGC
ATTCAAAATCTTGTCCAAGAACTCAAGTCACTTTACATCTATTAAGTCTTTGGAGACTTCATAA
TTTTTCTAGGAAAGGTGTTAGTGGTGTGTTTCACTGTTTTTGGAGGACTCATGGCTTTTAACTAC
AATCGGGCATTCCAGGTGTGGGCAGTCCCTCTGTTATTGGTAGCTTTTTTGCCTACTTAGTAGC
CCATAGTTTTTTATCTGTGTTTGAAACTGTGCTGGATGCACCTTTTCTGTGTTTGTGCTGTTGATC
TGGAACAAATGATGGATCGTCAGAAAAGCCCTACTTTATGGATCAAGAATTTCTGAGTTTCGTA
AAAAGGAGCAACAAATTAAACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAGGAATGAGGA
GGGAACAGAACTCCAGGCCATTGTGAGATAGATACCCATTTAGGTATCTGTACCTGGAACAACTT
TCCTTCTAAGAGCCATTTACAGAATAGAAGATGAGACCACTAGAGAAAAGTTAGTGAATTTTTTT
TTAAAAGACCTAATAAACCCCTATTCTTCCTCAAAA

FIGURE 32

LDTERENMKCVLGFAIVSTGITAVLLVLI FVL RKRIKLTVELFQITNKAISSAPFLLFQPLWTF
ILIFFWVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFILACQOMTIA
GAVVTCYFNRSKNPPDHPILSSLSILFFYHQGT VVKGSFLISVVRIPRIIVMYMQNALKEQQHG
ALSRYLFRCCYCCFWCLDKYLLHLNQAYTTTAINGTDFCTSAKDAFKILSKNSSHFTSINCFGD
FIIFLGKVLVVCFTVFGGLMAFNYNRAQVWAVPLLLVAFFAYLVAHSFLSVFETVLDALFLCFA
VDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRNEEGTELQAIVR

Important features:

Signal peptide:

amino acids 1-20

Putative transmembrane domains:

amino acids 35-54, 75-97, 126-146, 185-204, 333-350, 352-371

N-glycosylation sites.

amino acids 204-208, 295-299, 313-317

N-myristoylation sites.

amino acids 147-153, 178-184, 196-202, 296-275, 342-348

FIGURE 34

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKCPAG
CQDPKYHVYGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYSNGVQSLSLPRWR
ESFIVLESKPKKGVITYPSALTYSSSKSPAAQAGETTKAYQRPPIPGTTAQPVITLMQLLAVTVAVA
TPTTLPRPSPSAASTTSIPRPQSVGHRSEQEMDLWSTATYTSSQNRPRADPGIQRQDPGAAAFQKP
VGADVSLGLVPKEELSTQSLPEVSLGDPNCKIDLSFLIDGSTSIGKRRFRIQKQLLADVAQALDI
GPAGPLMGVVQYGDNPATFNLKTHNTSRDLKTAIEKITQRGGLSNVGRAISFVTKNFFSKANGN
RSGAPNVVVVMVDGWPTDKVEEASRLARESGINIFFITIEGAAENEKQYVVEPNFANKAVCRTNG
FYSLHVQSWFGLHKTLPVLRVCDTDLACSKTCLNSADIGFVIDGSSSVGTGNFRTVLQFVTN
LTKEFEISDITRIGAVQYTYEQRLFEGFDKYSSKPDILNAIKRVGYWSGGTSTGAAINFALEQL
FKKSKPNKRKLMILITDGRSYDDVRIPAMAAHLKGVITYAIGVAAWAAQEELEVIATHPARDHSFF
VDEFNLHQYVPRIIQNICTEFNSQPRN

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 181-200

N-glycosylation sites.

amino acids 390-394, 520-524

N-myristoylation sites.

amino acids 23-29, 93-99, 115-121, 262-268, 367-373, 389-395,
431-437, 466-472, 509-515, 570-576, 571-577, 575-581, 627-633

Amidation site.

amino acids 304-308

FIGURE 35

CCGAGCACAGGAGATTGCCTGCGTTTAGGAGGTGGCTGCGTTGTGGGAAAAGCTATCAAGGAAGAAATTGC
CAAACCATGTCTTTTTTCTGTTTCAGAGTAGTTCACAACAGATCTGAGTGTTTTAATTAAGCATGGAAT
ACAGAAAACAACAAAAAAGCTTAAGCTTTAATTTTCATCTGGAATTCACAGTTTTCTTAGCTCCCTGGACCC
GGTTGACCTGTTGGCTCTTCCCGCTGGCTGCTCTATCACGTGGTGTCTCCGACTACTACCCCGAGTGTA
AAGAACCCTTCGGCTCGCGTGCTTCTGAGCTGCTGTGGAATGGCCTCGGCTCTCTGGACTGTCTTCCGAGTA
GGATGTCACTGAGATCCCTCAAATGGAGCCTCCTGCTGCTGTCACTCCTGACTTTCTTTGTGATGTGGTAC
CTCAGCCTTCCCCACTACAATGTGATAGAACGCGTGAAGTGGATGTACTTCTATGAGTATGAGCCGATTTA
CAGACAAGACTTTCACTTCACACTTCGAGAGCATTCAAACTGCTCTCATCAAAATCCATTTCTGGTCATTC
TGGTGACCTCCACCCCTTCAGATGTGAAGCCAGGCAGGCCATTAGAGTTACTTGGGGTGAAAAAAGTCT
TGGTGGGGATATGAGGTTCTTACATTTTCTTATTAGGCCAAGAGGCTGAAAAGGAAGACAAAATGTTGGC
ATTGTCCTTAGAGGATGAACACCTTCTTTATGGTGACATAATCCGACAAGATTTTTTAGACACATAATAATA
ACCTGACCTTGAAGAACATTATGGCATTGAGGTGGGTAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT
AAGACAGACACTGATGTTTTCATCAATACTGGCAATTTAGTGAAGTATCTTTTAAACCTAAACCACTCAGA
GAAGTTTTTCACAGGTATCCTCTAATTGATAATTATTCCTATAGAGGATTTTACCAAAAAACCCATATTT
CTTACCAGGAGTATCCTTTCAAGGTGTTCCCTCCATACTGCAGTGGGTGGGTGGTGGTGGTGGTGGTGGT
TTGGTGCCCAAGGATCTATGAAATGATGGGTACGTAAGAACCCATCAAGTTTGAAGATGTTTATGTGGGAT
CTGTTTGAATTTATTAAGTGAACATTCATATCCAGAAGACACAAATCTTTTCTTCTATATAGAATCC
ATTTGGATGTCTGTCAACTCAGACGTGTGATTGCAGCCCATGGCTTTTCTTCCAAGGAGATCATCACTTTT
TGGCAGGTGATGCTAAGGAACACCACATGCCATTATTAAGTTCACATTCACAAAAAGCCTAGAAGGACAG
GATACCTTGTGGAAAGTGTAAATAAAGTAGGTACTGTGGAAAATTCATGGGGAGGTGAGTGTGTGGCTT
ACACTGAACTGAAACTCATGAAAAACCCAGACTGGAGACTGGAGGGTTACACTTGTGATTTATTAGTCAGG
CCCTTCAAAGATGATATGTGGAGGAATTAATATATAAGGAATTGGAGGTTTTTGGCTAAAGAAATTAATAGG
ACCAACAATTTGGACATGTCATTCTGTAGACTAGAATTTCTTAAAGGGTGTACTGAGTTATAAGCTCA
CTAGGCTGTAAAAACAAACAATGTAGAGTTTATTATTATGAACAATGTAGTCACTTGAAGGTTTTGTGTA
TATCTTATGTGGATTACCAATTTAAAAATATATGTAGTTCTGTGTCAAAAAACTTCTTCACTGAAGTTATA
CTGAACAAAATTTTACCTGTTTTTGGTCATTTATAAAGTACTTCAAGATGTTGCAGTATTTACAGTTATT
ATTATTTAAATTAATTTCACTTTTGTGTTTTTAAATGTTTTGACGATTTCATACAAGATAAAAGGATAG
TGAATCATTCTTTACATGCAACATTTCCAGTTACTTAACTGATCAGTTTATTATTGATACATCACTCCA
TTAATGTAAAGTCATAGGTCATTATTGCATATCAGTAATCTCTTGGACTTTGTAAATATTTTACTGTGGT
AATATAGAGAAGAATTAAAGCAAGAAATCTGAAAA

FIGURE 36

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPYRQDFHF
 TLREHSNCSHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSWWGYEVLTFLLGQEA EKEDKMLA
 LSLEDEHLLYGDIIRQDFLDTYNNLT LKTIMAFRWVTEFCPNAKYVMKTD TDVFINTGNLVKYL
 NLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQEYPFKVFPPYCSGLGYIMSRDLVPRIYEMMGHV
 KPIKFEDVYVGICLNLKVNIIHIPEDTNLFFLYRIHL DVCQLRRVIAAHGFSSKEIITFWQVMLR
 NTTCHY

Important features:

Type II transmembrane domain:

amino acids 20-39

N-glycosylation sites.

amino acids 72-76, 154-158, 198-202, 212-216, 326-330

Glycosaminoglycan attachment site.

amino acids 239-243

Ly-6 / u-PAR domain proteins.

amino acids 23-37

N-myristoylation site.

amino acids 271-277

FIGURE 37

FIGURE 38

MELGCWTQLGLTFLQLLLISSLPREYTVINEACPGAENIMCRECCEYDQIECVCPGKREVVGYT
 IPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGLDDFYVKGFYCAECRAGWYGGDCMRGQ
 VLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRDGDNRDGOII
 KRVCGNERPAPIQSIGSSSLHVLFSHSDGSKNFDGFAHYEEITACSSSPCFHDGTCVLDKAGSYKC
 ACLAGYTGQRCENLLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTVVSFFCNSYVLSGNE
 KRTCQQNGEWSGKQPICIKACREPKISDLVRRRVLPMPQVQSRETPLHQLYSAAFSKQKLQSAPTK
 KPALPFGDLPMGYQHLHTQLQYECISPFYRRLGSSRRTCLRTGKWSGRAPSCIPICGKIENITAP
 KTQGLRWPWQAAIYRRTSGVHDGSLHKGAWFLVCSGALVNERTVVVAHCVTDLGKVTMIKTADL
 KVVLGKFYRDDDRDEKTIQSLQISAILHFNYPILLDADIAILKLLDKARISTRVQPICLAASR
 DLSTSFQESHITVAGWNVLADVRSPGFKNDFLRSGVSVVDSLLCEEQHEDHGIPVSVTDNMFCA
 SWEPTAPSDICTAETGGIAAVSFPGRASPEPRWHLMGLVSWSYDKTCSHRLSTAFTKVLFPKDWI
 ERNMK

Important features of the protein:

Signal peptide:

amino acids 1-23

EGF-like domain cysteine pattern signature.

amino acids 260-272

N-glycosylation sites.

amino acids 96-100, 279-283, 316-320, 451-455, 614-618

N-myristoylation sites.

amino acids 35-41, 97-103, 256-262, 284-290, 298-304, 308-314,
 474-480, 491-497, 638-644, 666-672

Amidation site.

amino acids 56-60

Serine proteases, trypsin family.

amino acids 489-506

CUB domain proteins profile.

amino acids 150-167

FIGURE 39

GGTTCCTACATCCTCTCATCTGAGAATCAGAGAGCATAATCTTCTTACGGGCCCGTGATTATTAAACGTGGCTTAATC
 TGAAGGTTCTCAGTCAAATCTTTGTGATCTACTGATTGTGGGGGCATGGCAAGGTTTGCTTAAAGGAGCTTGGCTGG
 TTTGGGGCCCTTGTAGCTGACAGAAGGTGGCCAGGGAGAATGCAGCACACTGCTCGGAGAATGAAGGCGCTTCTGTTGC
 TGGTCTTGCCTTGGCTCAGTCTGTAACTACATTGACAATGTGGGCAACCTGCACTTCCTGTATTAGAACTCTGTA
 AAGGTGCCTCCCACTACGGCCTGACCAAAGATAGGAAGAGGCGCTCACAAGATGGCTGTCCAGACGGCTGTGCGAGCC
 TCACAGCCACGGCTCCCTCCCGAGAGGTTTCTGCAGCTGCCACCATCTCCTTAATGACAGACGAGCCTGGCCTAGACA
 ACCCTGCCTACGTGTCTCGGGCAGAGGACGGGCAGCCAGCAATCAGCCCAGTGGACTCTGGCCGGAGCAACCGAACTA
 GGGCACGGCCCTTTGAGAGATCCACTATTAGAAGCAGATCATTTAAAAAATAAATCGAGCTTTGAGTGTCTTCGAA
 GGACAAAGAGCGGGAGTGCAGTTGCCAACCATGCCGACCAGGGCAGGGAAAATTCTGAAAACACCACTGCCCTGAAG
 TCTTTCCAAGGTTGTACCACCTGATTCCAGATGGTGAAATTACCAGCATCAAGATCAATCGAGTAGATCCAGTGAAA
 GCCTCTCTATTAGGCTGGTGGGAGGTAGCGAAACCCCACTGGTCCATATCATTTCAACACATTTATCGTGATGGGG
 TGATCGCCAGAGACGGCCGGCTACTGCCAGGAGACATCATTTAAAGGTCAACGGGATGGACATCAGCAATGTCCCTC
 ACAACTACGCTGTGCGTCTCTGCGGCAGCCCTGCCAGGTGCTGTGGCTGACTGTGATGCGTGAACAGAAGTCCGCA
 GCAGGAACAATGGACAGGCCCGGATGCCACAGACCCGAGATGACAGCTTTCATGTGATTCTCAACAAAAGTAGCC
 CCGAGGAGCAGCTTGAATAAACTGGTGGCAGAGGTGGATGAGCCTGGGGTTTCATCTTCAATGTGTGGATGGCG
 GTGTGGCATATCGACATGGTCAGCTTGAGGAGAATGACCGTGTGTAGCCATCAATGGACATGATCTTCGATATGGCA
 GCCCAGAAAGTGGCGCTCATCTGATTCAAGCCAGTGAAAGACGTGTTACCTCGTGTGTCCCGCCAGGTTCCGGCAGC
 GGAGCCCTGACATCTTTCAGGAAGCCGGCTGGAACAGCAATGGCAGCTGTTCCCGAGGGCCAGGGGAGAGGAGCAACA
 CTCCCAAGCCCTCCATCCTACAATTACTTGTCTATGAGAAGGTGGTAAATATCCAAAAGACCCCGGTGAATCTCTCG
 GCATGACCGTCGCAGGGGAGCATCACATAGAGAATGGGATTTGCCTATCTATGTCTCATGTTGAGCCCGGAGGAG
 TCATAAGCAGAGATGGAAGAATAAAACAGGTGACATTTTGTGAATGTGGATGGGGTCGAACGACAGAGGTGAGCC
 GGAGTGAGGCAGTGGCATTATTGAAAAGAACATCATCTCGATAGTACTCAAAGCTTTGGAAGTCAAAGAGTATGAGC
 CCCAGGAAGACTGCAGCAGCCAGCAGCCCTGGACTCCAACCACAACATGGCCCCACCCAGTGACTGGTCCCATCCT
 GGGTCATGTGGCTGGAATTACCACGGTGCTTGTATAACTGTAAAGATATTGTATTACGAAGAAACACAGCTGGAAGTC
 TGGGCTTCTGCATTGTAGGAGGTTATGAAGAATACAATGGAACAAACCTTTTTCATCAAATCCATTGTTGAAGGAA
 CACCAGCATACAATGATGGAAGAATTAGATGTGGTGATATTCTTCTGCTGTCAATGGTAGAAGTACATCAGGAATGA
 TACATGCTTGCTTGGCAAGACTGCTGAAAGAATTAAAGGAAGAATTACTCTAACTATTGTTTCTTGGCCTGGCACTT
 TTTTATAGAAATCAATGATGGGTCAGAGGAAAACAGAAAAATCACAAATAGGCTAAGAAGTTGAAACACTATATTATC
 TTGTGAGTTTTTATATTAAAGAAAAGAAATACATTGTAAAAATGTCAGGAAAAGTATGATCATCTAATGAAAGCCAGTT
 ACACCTCAGAAAATATGATCCAAAAAATAAACTACTAGTTTTTTTTTTCAGTGTGGAGGATTTCTCATTACTCTAC
 AACATTGTTTTATATTTTTCTATTCAATAAAAAGCCCTAAACAACATAAATGATTGATTTGTATACCCCACTGAATT
 CAAGCTGATTTAAATTTAAATTTGGTATATGCTGAAGTCTGCCAAGGGTACATTATGGCCATTTTAAATTTACAGCT
 AAAATATTTTTTAAATGCATTGCTGAGAAACGTTGCTTTCATCAAACAAGAATAAATATTTTTTCAGAAGTTAAA

FIGURE 40

MKALLLLVLPWLSPANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPCDGCASLTATAPS
 PEVSAAATISLMTDEPGLDNPAYVSSAEDGQPAISPVDSGRSNRTRARPFERSTIRSRSFKKINR
 ALSVLRRTKSGSAVANHADQGRESENTTAPEVFPRLYHLIPDGEITSIKINRVDPSESLSIRLV
 GGSETPLVHII IQHIYRDGVIARDGRLLPGDII LKVNMGDISNVPHNYAVRLLRQPCQVLWLTVM
 REQKFRSRNNGQAPDAYRPRDDSFHVILNKSSPEEQLGIKLVRKVDEPGVFIFNVLDGGVAYRHG
 QLEENDRVLAINGHDLRYGSPESAHLIQASERRVHLVVSQRQRSPDIFQEAGWNSNGSWSPG
 PGRSNTPKPLHPTITCHEKVNIQKDPGESLGMTVAGGASHREWDLPIYVISVEPGGVISRDGR
 IKTGDI LLNVVGVELTEVSRSEAVALLKRTSSSIVLKALEVKEYEQEDCSSPAALDSNHNMAP
 SDWSPSWVMWLELPRCLYNCKDIVLRRNTAGSLGFCIVGGYEEYNGNKPFFIKSIVEGTPAYNDG
 RIRCGDILLAVNGRSTSGMIHACLARLLKELKGRITLTIVSWPGTFL

Important features:

Signal peptide:

amino acids 1-15

N-glycosylation sites.

amino acids 108-112, 157-161, 289-293, 384-388

Tyrosine kinase phosphorylation sites.

amino acids 433-441, 492-500

N-myristoylation sites.

amino acids 51-57, 141-147, 233-239, 344-350, 423-429, 447-453,
 467-473, 603-609

FIGURE 42

MGFNLT FHLSYKFRLLLLLTLCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTLILGKGKTLTN
EASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRVAILVPH
RNREKHLMYLLEHLHPFLQRQQLDYGIYVIHQAEKKFNRAKLLNVGYLEALKEENWDCFI FHDV
DLVPENDFNLYKCEEHPKHLVVGRNSTGYRLRYSGYFGGV TALSREQFFKVNGFSNNYWGWGGE
DDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSRVWRTDGLSSCSYKLV
SVEHNPLYINITVDFWFGA

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 4-8, 220-224, 335-339

Xylose isomerase proteins.

amino acids 191-202

FIGURE 43

GCTCAAGACCCAGCAGTGGGACAGCCAGACAGACGGCACGATGGCACTGAGCTCCCAGATCTGGG
CCGCTTGCCTCCTGCTCCTCCTCCTCCTCGCCAGCCTGACCAGTGGCTCTGTTTTCCCACAACAG
ACGGGACAACCTGCAGAGCTGCAACCCCAGGACAGAGCTGGAGCCAGGGCCAGCTGGATGCCCAT
GTTCCAGAGGGGAAGGAGGCGAGACACCCACTTCCCCATCTGCATTTTCTGCTGCGGCTGCTGTC
ATCGATCAAAGTGTGGGATGTGCTGCAAGACGTAGAACCTACCTGCCCTGCCCCCGTCCCCTCCC
TTCCTTATTTATTCCTGCTGCCCCAGAACATAGGTCTTGAATAAAATGGCTGGTTCTTTGTTT
TCCAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 44

MALSSQIWAACLLLLLLLLASLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRRDTHFPI
CIFCCGCCHRSKCGMCCKT

Important features:

Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 58-59

N-myristoylation site.

amino acids 44-50

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 1-12

FIGURE 45

GTGGCTTCATTTTCAGTGGCTGACTTCCAGAGAGCAATATGGCTGGTTCCCCAACATGCCTCACCC
 TCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTGGTCGGT
 TCCGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTCCAAAGTAAAGCAAGTTGACTCTATTGTCTG
 GACCTTCAACACAACCCCTCTTGTCCACATACAGCCAGAAGGGGGCACTATCATAGTGACCCAAA
 ATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAGCTCAGCAAACCTGAAG
 AAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCACTCCAGCAGCCCTCCACCCA
 GGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCAGAGCA
 ATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATGGAACATGGGGAAGAGGATGTGATT
 TATACCTGGAAGGCCCTGGGGCAAGCAGCCAATGAGTCCCATAATGGGTCCATCCTCCCCATCTC
 CTGGAGATGGGGAGAAAGTGATATGACCTTCATCTGCGTTGCCAGGAACCTGTGAGCAGAACT
 TCTCAAGCCCCATCCTTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCTCCATG
 GTCCTCCTGTGTCTCCTGTTGGTGCCCTCCTGCTCAGTCTCTTTGTACTGGGGCTATTTCTTTG
 GTTTCTGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGAGTGGACATTTGTCGGG
 AAACCTCCTAACATATGCCCCATTCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAAT
 AGAACAATCCTAAAGGAAGATCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAAAGAT
 GGAAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTTGCCTATGAGAATGTTA
 TCTTAGACAGCAGTGCACTCCCCTAAGTCTCTGCTCA

[illegible]

Important features:

amino acids 1-22

amino acids 224-250

amino acids 229-251

amino acids 98-102, 142-146, 148-152, 172-176, 176-180, 204-208,
291-295

FIGURE 47

GGCTCGAGCGTTTCTGAGCCAGGGGTGACCATGACCTGCTGCGAAGGATGGACATCCTGCAATGG
ATTCAGCCTGCTGGTTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATACCTCTAATTGTCAGCT
TAGTTGAGGAAGACCAATTTTCTCAAACCCCATCTCTTGCTTTGAGTGCGTGGTCCCAGGAATT
ATAGGAGCAGGTCTGATGGCCATTCCAGCAACAACAATGTCCTTGACAGCAAGAAAAAGAGCGTG
CTGCAACAACAGAACTGGAATGTTTCTTTCATCATTTTTTCAGTGTGATCACAGTCATTGGTGCTC
TGTATTGCATGCTGATATCCATCCAGGCTCTCTTAAAAGGTCTCTCATGTGTAATTCTCCAAGC
AACAGTAATGCCAATTGTGAATTTTCATTGAAAAACATCAGTGACATTCATCCAGAATCCTTCAA
CTTGTCAGTGGTTTTTCAATGACTCTTGTCACCTCCTACTGGTTTCAATAAACCCACCAGTAACG
ACACCATGGCGAGTGGCTGGAGAGCATCTAGTTTCCACTTCGATTCTGAAGAAAACAAACATAGG
CTTATCCACTTCTCAGTATTTTtaggtctattgcttggtggaattctggaggtcctgtttgggct
CAGTCAGATAGTCATCGGTTTCCTTGGCTGTCTGTGTGGAGTCTCTAAGCGAAGAAGTCAAATTG
TGTAGTTTAATGGGAATAAAATGTAAGTATCAGTAGTTTGAAAAAAAAAA

FIGURE 48

MTCCEGWTSCNGFSLLVLLLLGVVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMAIPA
 TTMSLTARKKRACCNRTGMFLSSFFSVITVIGALYCM LISIQALLKGPLMCNSPSNSNANCEFSL
 KNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDSEENKHRLIHFSVFLGL
 LLVGILEVLFLGLSQIVIGFLGCLCGVSKRRSQIV

Important features:

Transmembrane domains:

amino acids 10-31 (type II), 50-72, 87-110, 191-213

N-glycosylation sites.

amino acids 80-84, 132-136, 148-152, 163-167

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 223-227

N-myristoylation sites.

amino acids 22-28, 54-60, 83-89, 97-103, 216-222

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 207-218

TNFR/NGFR family cysteine-rich region protein.

amino acids 4-12

FIGURE 49

ATCCGTTCTCTGCGCTGCCAGCTCAGGTGAGCCCTCGCCAAGGTGACCTCGCAGGACACTGGTGA
 AGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTGGAGCAG
 ATCCGTGGGCTGCAGACCCCCGCCCCAGTGCCTCTCCCCCTGCAGCCCTGCCCCCTCGAACTGTGA
CATGGGAGAGAGTGACCCTGGCCCTTCTCCTACTGGCAGGCCTGACTGCCTTGAAGCCAATGACC
 CATTTGCCAATAAAGACGATCCCTTCTACTATGACTGGAAAAACCTGCAGCTGAGCGGACTGATC
 TGCGGAGGGCTCCTGGCCATTGCTGGGATCGCGGCAGTTCTGAGTGGCAAATGCAAATACAAGAG
 CAGCCAGAAGCAGCACAGTCCTGTACCTGAGAAGGCCATCCCACTCATCACTCCAGGCTCTGCCA
 CTACTTGCTGAGCACAGGACTGGCCTCCAGGGATGGCCTGAAGCCTAACACTGGCCCCCAGCACC
 TCCTCCCCTGGGAGGCCTTATCCTCAAGGAAGGACTTCTCTCCAAGGGCAGGCTGTTAGGCCCCCT
 TTCTGATCAGGAGGCTTCTTTATGAATTAAACTCGCCCCACCACCCCTCA

FIGURE 50

MERVTLALLLLAGLTALEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAGIAAVLSGKCKYKS
SQKQHSPVPEKAIPILITPGSATTC

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 36-59

N-myristoylation sites.

amino acids 41-47, 45-51, 84-90

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7.

amino acids 54-67

FIGURE 51

GTGGACTCTGAGAAGCCCAGGCAGTTGAGGACAGGAGAGAGAAGGCTGCAGACCCAGAGGGAGGG
 AGGACAGGGAGTCGGAAGGAGGAGGACAGAGGAGGGCACAGAGACGCAGAGCAAGGGCGGCAAGG
 AGGAGACCCTGGTGGGAGGAAGACACTCTGGAGAGAGAGGGGGCTGGGCAGAGATGAAGTTCCAG
 GGGCCCCCTGGCCTGCCTCCTGCTGGCCCTCTGCCTGGGCAGTGGGGAGGCTGGCCCCCTGCAGAG
 CGGAGAGGAAAGCACTGGGACAAATATTGGGGAGGCCCTTGGACATGGCCTGGGAGACGCCCTGA
 CGGAAGGGGTGGGAAAGGCCATTGGCAAAGAGGCCGAGGGGCAGCTGGCTCTAAAGTCAGTGAG
 GCCCTTGGCCAAGGGACCAGAGAAGCAGTTGGCACTGGAGTCAGGCAGGTTCCAGGCTTTGGCGC
 AGCAGATGCTTTGGGCAACAGGGTCGGGGAAGCAGCCCATGCTCTGGGAAACACTGGGCACGAGA
 TTGGCAGACAGGCAGAAGATGTCATTGACACGGAGCAGATGCTGTCCGCGGCTCCTGGCAGGGG
 GTGCTTGGCCACAGTGGTGCTTGGGAACTTCTGGAGGCCATGGCATCTTTGGCTCTCAAGGTGG
 CCTTGGAGGCCAGGGCCAGGGCAATCCTGGAGGTCTGGGGACTCCGTGGGTCCACGGATAACCCG
 GAAACTCAGCAGGCAGCTTTGGAATGAATCCTCAGGGAGCTCCCTGGGGTCAAGGAGGCAATGGA
 GGGCCACCAAACTTTGGGACCAACACTCAGGGAGCTGTGGCCAGCCTGGCTATGGTTCAAGTGAG
 AGCCAGCAACCAGAATGAAGGGTGCACGAATCCCCACCATCTGGCTCAGGTGGAGGCTCCAGCA
 ACTCTGGGGGAGGCAGCGGCTCACAGTCGGGCAGCAGTGGCAGTGGCAGCAATGGTGACAACAAC
 AATGGCAGCAGCAGTGGTGGCAGCAGCAGTGGCAGCAGCAGTGGCAGCAGCAGTGGCGGCAGCAG
 TGGCGGCAGCAGTGGTGGCAGCAGTGGCAACAGTGGTGGCAGCAGAGGTGACAGCGGCAGTGAGT
 CCTCCTGGGGATCCAGCACCGGCTCCTCCTCCGGCAACCACGGTGGGAGCGGCGGAGGAAATGGA
 CATAAACC CGGTGTGAAAAGCCAGGGAATGAAGCCCGGGAGCGGGGAATCTGGGATTCAGGG
 CTT CAGAGGACAGGGAGTTTCCAGCAACATGAGGGAATAAGCAAAGAGGGCAATCGCCTCCTTG
 GAGGCTCTGGAGACAATTATCGGGGGCAAGGGTCGAGCTGGGGCAGTGGAGGAGGTGACGCTGTT
 GGTGGAGTCAATACTGTGAACCTCTGAGACGTCCTGGGATGTTTAACTTTGACACTTTCTGGAA
 GAATTTTAAATCCAAGCTGGGTTTCATCAACTGGGATGCCATAAACAAGGACCAGAGAAGCTCTC
 GCATCCCGTGACCTCCAGACAAGGAGCCACCAGATTGGATGGGAGCCCCCACACTCCCTCCTTAA
 AACACCACCTCTCATCACTAATCTCAGCCCTTGCCCTTGAAATAAACCTTAGCTGCCCCACAAA
 AA
 AA

FIGURE 52

MKFQGPLACLLALCLGSGEAGPLQSGEESTGTNIGEALGHGLGDALSEGKGAIGKEAGGAAGSKVS
EALGQGTREAVGTGVRQVPFGGAADALGNRVGEAAHALGNTGHEIGRQAEDVIRHGADAVRGSWQGV
GHSGAWETSGGHGIFGSQGGGLGGQGGNPGGLGTPWVHGYPGNSAGSFGMNPQGAPWGQGGNGGPPNF
GTNTQGAVAQPGYGSVRASNQNEGCTNPPPSGSGGSSNSGGGSGSQSGSSGSGSNGDNNNGSSSGGS
SSGSSSGSSSGGSSGGSSGGSSGNSGSGSRGDSGSESSWGSSTGSSSGNHGGSGGGNGHKPGCEKPGNE
ARGSGESGIQGFRRQGVSSNMREISKEGNRLLGGSGDNYRGQSSWGSGGGDAVGGVNTVNSETPGM
FNFDTFWKNFKSKLGFINWDAINKDQRSSRIP

Signal peptide:

amino acids 1-21

N-glycosylation site.

amino acids 265-269

Glycosaminoglycan attachment site.

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

Casein kinase II phosphorylation site.

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

N-myristoylation site.

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70, 74-80,
90-96, 96-102, 130-136, 140-146, 149-155, 152-158, 155-161,
159-165, 163-169, 178-184, 190-196, 194-200, 199-205, 218-224,
236-242, 238-244, 239-245, 240-246, 245-251, 246-252, 249-252,
253-259, 256-262, 266-272, 270-276, 271-277, 275-281, 279-285,
283-289, 284-290, 287-293, 288-294, 291-297, 292-298, 295-301,
298-304, 305-311, 311-317, 315-321, 319-325, 322-328, 323-329,
325-331, 343-349, 354-360, 356-362, 374-380, 381-387, 383-389,
387-393, 389-395, 395-401

Cell attachment sequence.

amino acids 301-304

FIGURE 53

GGAGAAGAGGTTGTGTGGGACAAGCTGCTCCCGACAGAAGGATGTCGCTGCTGAGCCTGCCCTGG
CTGGGCCTCAGACCGGTGGCAATGTCCCATGGCTACTCCTGCTGCTGGTTGTGGGCTCCTGGCT
ACTCGCCCGCATCCTGGCTTGGACCTATGCCCTTCTATAACAACTGCCGCCGGCTCCAGTGTTC
CACAGCCCCAAAACGGAACCTGGTTTTGGGGTCACCTGGGCCTGATCACTCCTACAGAGGAGGGC
TTGAAGGACTCGACCCAGATGTCGGCCACCTATCCCAGGGCTTTACGGTATGGCTGGGTCCCAT
CATCCCCCTTCATCGTTTTATGCCACCCTGACACCATCCGGTCTATCACCAATGCCTCAGCTGCCA
TTGCACCCAAGGATAATCTCTTCATCAGGTTCTGAAGCCCTGGCTGGGAGAAGGGATACTGCTG
AGTGGCGGTGACAAGTGGAGCCGCCACCGTCGGATGCTGACGCCCGCCTCCATTTCAACATCCT
GAAGTCCTATATAACGATCTTCAACAAGAGTGCAAACATCATGCTTGACAAGTGGCAGCACCTGG
CCTCAGAGGGCAGCAGTCGCTGACATGTTTGAGCACATCAGCCTCATGACCTTGGACAGTCTA
CAGAAATGCATCTTCAGCTTTGACAGCCATTGTGAGGAGAGGCCAGTGAATATATTGCCACCAT
CTTGGAGCTCAGTGCCCTTGTAGAGAAAAGAAGCCAGCATATCCTCCAGCACATGGACTTTCTGT
ATTACCTCTCCCATGACGGGCGCGCTTCCACAGGGCCTGCCGCTGGTGCATGACTTCACAGAC
GCTGTCATCCGGGAGCGGCGTCGCACCCTCCCCACTCAGGGTATTGATGATTTTTTCAAAGACAA
AGCCAAGTCCAAGACTTTGGATTTCATTGATGTGCTTCTGCTGAGCAAGGATGAAGATGGGAAGG
CATTGTGATGAGGATATAAGAGCAGAGGCTGACACCTTCATGTTTGGAGGCCATGACACCACG
GCCAGTGGCCTCTCCTGGGTCTGTACAACCTTGCGAGGCACCCAGAATACCAGGAGCGCTGCCG
ACAGGAGGTGCAAGAGCTTCTGAAGGACCGCGATCCTAAAGAGATTGAATGGGACGACCTGGCCC
AGCTGCCCTTCTGACCATGTGCGTGAAGGAGAGCCTGAGGTTACATCCCCAGCTCCCTTCATC
TCCCGATGCTGCACCCAGGACATTGTTCTCCAGATGGCCGAGTCATCCCCAAAGGCATTACCTG
CCTCATCGATATTATAGGGGTCCATCACAACCCAACTGTGTGGCCGGATCCTGAGGTCTACGACC
CCTTCCGCTTTGACCCAGAGAACAGCAAGGGGAGGTACCTCTGGCTTTTATTCCTTTCTCCGCA
GGGCCAGGAACATGCATCGGGCAGGCGTTCGCCATGGCGGAGATGAAAGTGGTCTGCGGTTGAT
GCTGCTGCACTTCGGTTTCTGCCAGACCACACTGAGCCCCGCAGGAAGCTGGAATTGATCATGC
GCGCCGAGGGCGGGCTTTGGCTGCGGGTGGAGCCCCTGAATGTAGGCTTGACAGTACTTTCTGAC
CCATCCACCTGTTTTTTTGCAGATTGTCATGAATAAACGGTGTGTCAA

FIGURE 54

MSLLSLPWLGLRPVAMSPWLLLLLVVGSWLLARILAWTYAFYNNCRRLQCFPPKRNWFWGHLG
 LITPTEGLKDSTQMSATYSQGFTVWLGPIIPFIVLCHPDITRSITNASAAIAPKDNLFIRFLKP
 WLGEIGILLSGGDKWSRHRMLTPAFHFNILKSYITIFNKSANIMLDKWQHLASEGSSRLDMFEHI
 SLMTLDSLQKCIFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYYLSHDGRRFHRAC
 RLVHDFDPAVIRERRRTLPTQGIDDFKDKAKSKTLDLFDVLLLSKDEDGKALSDEDIRAEADTF
 MFGGHDTTASGLSWVLYNLARHPEYQERCQEVQELLKDRDPKEIEWDDLAQLPFLTMCVKESLR
 LHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVWPDPEVYDPFRFDPENSKGRSP
 LAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHTEPRRKLELIMRAEGGLWLRVEPLN
 VGLQ

Important features:

Transmembrane domains:

amino acids 13-32 (type II), 77-102

Cytochrome P450 cysteine heme-iron ligand signature.

amino acids 461-471

N-glycosylation sites.

amino acids 112-116, 168-172

FIGURE 55

ATCGCATCAATTGGGAGTACCATCTTCCTCATGGGACCAGTGAAACAGCTGAAGCGAATGTTTGA
 GCCTACTCGTTTGATTGCAACTATCATGGTGCTGTTGTGTTTTGCACTTACCCTGTGTTCTGCCT
 TTTGGTGGCATAACAAGGGACTTGCACTTATCTTCTGCATTTTGCAGTCTTTGGCATTGACGTGG
 TACAGCCTTTCCTTCATACCATTGCAAGGGATGCTGTGAAGAAGTGTTTTGCCGTGTGTCTTGC
ATAATCATGGCCAGTTTTATGAAGCTTTGGAAGGCACTATGGACAGAAGCTGGTGGACAGTTTT
 GTAACATCTTCGAAACCTCTGTCTTACAGACATGTGCCTTTTATCTTGCAGCAATGTGTTGCTT
 GTGATTCGAACATTTGAGGGTTACTTTTGAAGCAACAATACATTCTCGAACCTGAATGTCAGTA
 GCACAGGATGAGAAGTGGGTTCTGTATCTTGTGGAGTGGAATCTTCCTCATGTACCTGTTTCCCTC
 TCTGGATGTTGTCCCACTGAATTCCCATGAATACAAACCTATTCAGCAACAGCAAAAAAAAAAAAA
 AA

Important features:

Signal peptide:

amino acids 1-33

Type II fibronectin collagen-binding domain protein.

amino acids 30-72

FIGURE 57

CGGCTCGAGCTCGAGCCGAATCGGCTCGAGGGGAGTGGAGCACCAGCAGGCCGCCAACATGCTCTGTCTGTGCCTG
TACGTGCCGGTCATCGGGGAAGCCAGACCGAGTTCAGTACTTTGAGTCGAAGGGGCTCCCTGCCGAGCTGAAGTCC
ATTTTCAAGCTCAGTGTCTTCATCCCTCCAGGAATTTCTCCACCTACCGCCAGTGGAGCAGAAAATTTGACAAAGCT
GGAGATAAGGACCTTGATGGGCAGCTAGACTTTGAAGAATTTGTCCATTATCTCCAAGATCATGAGAAGAAGCTGAGG
CTGGTGTAAAGATTTGGACAAAAAGATGATGGACGCATTGACGCGCAGGAGATCATGCAGTCCCTGCGGGACTTG
GGAGTCAAGATATCTGAACAGCAGGCAGAAAAAATTTCTCAAGAGCATGGATAAAACGGCAGATGACCATCGACTGG
AACGAGTGGAGAGACTACCACCTCCTCCACCCCGTGGAAAAACATCCCGAGATCATCCTCTACTGGAAGCATTCCACG
ATCTTTGATGTGGGTGAGAATCTAACGGTCCCGGATGAGTTCACAGTGGAGGAGAGGCAGACGGGGATGTGGTGAGA
CACCTGGTGGCAGGAGTGGGGCAGGGGCGTATCCAGAACCTGCACGGCCCCCTGGACAGGCTCAAGGTGCTCATG
CAGGTCCATGCTCCCGCAGCAACAACATGGGCATCGTTGGTGGCTTCACTCAGATGATTCGAGAAGGAGGGGCCAGG
TCACTCTGGCGGGGCAATGGCATCAACGTCTCAAATTTGCCCGGAATCAGCCATCAAATTCATGGCCTATGAGCAG
ATCAAGCGCCTTGTGGTAGTGACAGGAGACTCTGAGGATTCAGAGAGGGCTTGTGGCAGGGTCTTGGCAGGGGCC
ATCGCCAGAGCAGCATCTACCCAATGGAGGTCTGAAGACCCGATGGCGCTGCGGAAGACAGGCCAGTACTCAGGA
ATGCTGGAGTGGCGCAGGAGGATCTGGCCAGAGAGGGGGTGGCCGCTTCTACAAAGGCTATGTCCCAACATGCTG
GGCATCATCCCTATGCCGGCATCGACCTTGCAGTCTACGAGACGCTCAAGAATGCTGGCTGCAGCACTATGCAGTG
AACAGCGCGGACCCCGCGTGTGGTCTCTGGCTCTGGCCATGTCAGTACCTGTGGCAGCTGGCCAGCTAC
CCCTGGCCCTAGTCAGGACCCGGATGCAGGCGCAAGCCTTATTGAGGGCGCTCCGGAGGTGACCATGAGCAGCCTC
TTCAAACATATCTGCGGACCGAGGGGGCTTCGGGCTGTACAGGGGGCTGGCCCCCAACTTCATGAAGGTGATCCCA
GCTGTGAGCATCAGTACGTGGTCTACGAGAACCTGAAGATCACCTGGGCGTGCAGTGGCGGTGACGGGGAGGGC
CGCCCGGAGTGGAGTCTGATCTTGGGCGCAGCCTGGGCTGTGCAGCCATCTCATTTCTGTGAATGTGCCAACACT
AAGCTGTCTCGAGCCAAGCTGTGAAAACCTTAGACGCACCCGAGGGAGGGTGGGAGAGCTGGCAGGCCAGGGCTT
GTCCTGTGACCCAGCAGACCTCTGTGGTTCAGCGAAGACCAGGCATTCTTAGGGTCCAGGGTCAAGAGG
CTCCGGGCTCACATGTGTAAGGACAGGACATTTCTGCAGTGCCTGCCAATAGTGAGCTTGGAGCCTGGAGGCCGGCT
TAGTTCTTCCATTTACCCCTTGCAGCCAGCTGTTGGCCACGGCCCTGCCCTCTGGTCTGCCGTGCATCTCCCTGTGC
CCTCTTGTGCTGCTGCTGTCTGTGAGGTAAAGTGGGAGGAGGGCTACAGCCACATCCCAACCCCTCGTCCAATCCC
ATAATCATGATGAAAGGTGAGGTACGTGGCCTCCAGGCCTGACTTCCCAACCTACAGCATTGACGCCAATTTGGC
TGTGAAGGAAGAGGAAAGGATCTGGCCTTGTGGTCACTGGCATCTGAGCCCTGCTGATGGCTGGGGCTCTCGGGCATG
CTTGGGAGTGCAGGGGGCTCGGGCTGCCTGGCCTGGCTGCACAGAAGGCAAGTGTGGGGCTCATGGTCTCTGAGCT
GGCCTGGACCTGTGAGGATGGGCCCACTCAGAACCAACTCACTGTCCCACTGTGGCATGAGGGCAGTGGAGCA
CCATGTTTGAAGGCGAAGGGCAGAGCGTTTGTGTCTGGGGAGGGAAGGAAAGGTGTTGGAGGCCCTAATTTATGG
ACTGTTGGGAAAAGGGTTTGTCCAGAAGGACAAGCCGGACAAATGAGCGACTTCTGTGCTTCCAGAGGAAGACGAGG
GAGCAGGAGCTTGGCTGACTGCTCAGAGTCTGTTCTGACGCCCTGGGGTCTCTGTCCAACCCAGCAGGGGCGCAGC
GGGACAGCCCCACATTCACCTTGTGTCACTGCTTGGAACCTATTTATTTTGTATTTATTTGAACAGAGTTATGTCCT
AACTATTTTATAGATTTGTTTAATTAATAGCTTGTCAATTTTCAAGTTCAATTTTATTCATATTTATGTTTATGGTT
GATTGTACCTTCCCAAGCCCGCCAGTGGGATGGGAGGAGGAGGAGAAGGGGGCTTGGGCCGCTGCAGTCACATCT
GTCCAGAGAAATTCCTTTTGGACTGGAGGCAGAAAAAGCGGCCAGAGGCAGCAGCCCTGGCTCCTTTCTTGGCAG
GTTGGGGAAGGGCTTGGCCCGAGCCTTAGGATTTAGGGTTTGAAGTGGGCGTGGAGAGAGAGGGAGGAACCTCAAT
AACCTTGAAGGTGGAATCCAGTTATTTCTGCGCTGCGAGGGTTTCTTATTTCACTCTTTCTGAATGTCAAGGCAG
TGAGGTGCTCTCACTGTGAATTTGTGGTGGGCGGGGCTGGAGGAGAGGGTGGGGGCTGGCTCCCTCCCTCCAGC
CTTCTGCTGCCCTTGTAAATGCGCGCAACTGGCGACCTCAGGTTGCACTTCCATTCCACCAGAATGACCTGA
TGAGGAAATCTTCAATAGGATGCAAGATCAATGCAAAAAATTTGTATATATGAACATATAACTGGAGTCTGCAAAAAA
CAATTAAGAAAGAAATGGACGTTAGAAGTTGTCAATTAAGCAGCCTTCTAATAAGTTGTTTCAAGCTGAAAAAA
AA

FIGURE 58

MLCLCLYVPVIGEAQTEFQYFESKGLPAELKSI FKL SVFIPSQEFSTYRQWKQKIVQAGDKDL DG
 QLDFEEFVHYLQDHEKKLRLVFKILDKN DGRIDAQEIMQSLRDLGVKISEQQA EKILKSMDKNG
 TMTIDWNEW RDYHLLHPVENIPEI ILYWKHSTIFDVGENLTVPDEFTVEERQTGMWWRHLVAGGG
 AGAVSRTCTAPLDR LKVL MQVHASRSNNMGIVGGFTQMIREGGARSLWRGNGINVLKIAPESA IK
 FMAYEQIKRLVGSDQETLRIHERLVAGSLAGAI AQSSIYPMEVLKTRMALRKTGQYSGMLDCARR
 ILAREGVAAFYKGYVPNMLGIIPYAGIDLAVYETLKN AWLQHYAVNSADPGVFVLLACGTMSSTC
 GQLASYPLALVRTRMQAQASIEGAPEVTMSSLFKHILRTEGAFGLYRGLAPNFMKVIPAVSISYV
 VYENLKITLGVQSR

Important features:

Signal peptide:

amino acids 1-16

Putative transmembrane domains:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation sites.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

FIGURE 59

GGAAGGCAGCGGCAGCTCCACTCAGCCAGTACCCAGATACGCTGGGAACCTTCCCAGCCATGGC
TTCCCTGGGGCAGATCCTCTTCTGGAGCATAATTAGCATCATCATTATTCTGGCTGGAGCAATTG
CACTCATCATTGGCTTTGGTATTTAGGGAGACACTCCATCACAGTCACTACTGTGCGCTCAGCT
GGGAACATTGGGGAGGATGGAATCCTGAGCTGCACTTTTGAACCTGACATCAAACCTTTCTGATAT
CGTGATACAATGGCTGAAGGAAGGTGTTTTAGGCTTGGTCCATGAGTTCAAAGAAGGCAAAGATG
AGCTGTGCGAGCAGGATGAAATGTTAGAGGCCGGACAGCAGTGTGCTGATCAAGTGATAGTT
GGCAATGCCTCTTTGCGGCTGAAAAACGTGCAACTCACAGATGCTGGCACCTACAAATGTTATAT
CATCACTTCTAAAGCAAGGGGAATGCTAACCTTGAGTATAAACTGGAGCCTTCAGCATGCCGG
AAGTGAATGTGGACTATAATGCCAGCTCAGAGACCTTGCGGTGTGAGGCTCCCCGATGGTTCCCC
CAGCCCACAGTGGTCTGGGCATCCCAAGTTGACCAGGGAGCCAACTTCTCGGAAGTCTCCAATAC
CAGCTTTGAGCTGAACTCTGAGAAATGTGACCATGAAGGTTGTGTCTGTGCTCTACAATGTTACGA
TCAACAACACATACTCCTGTATGATTGAAAATGACATTGCCAAAGCAACAGGGGATATCAAAGTG
ACAGAAATCGGAGATCAAAGGCGGAGTCACTACAGCTGCTAAACTCAAAGGCTTCTCTGTGTGT
CTCTTCTTTCTTTGCCATCAGCTGGGCACCTTCTGCCTCTCAGCCCTTACCTGATGCTAAAAATAAT
GTGCCCTTGGCCACAAAAAAGCATGCAAAGTCATTGTTACAACAGGGATCTACAGAACTATTTTCAC
CACCAGATATGACCTAGTTTTATATTTCTGGGAGGAAATGAATTCATATCTAGAAGTCTGGAGTG
AGCAACAAGAGCAAGAAACAAAAAGAAGCCAAAGCAGAAGGCTCCAATATGAACAAGATAAAT
CTATCTTCAAAGACATATTAGAAGTTGGGAAAATAATTCATGTGAACTAGACAAGTGTGTTAAGA
GTGATAAGTAAATGCACGTGGAGACAAGTGCATCCCCAGATCTCAGGGACCTCCCCCTGCCTGT
CACCTGGGGAGTGAGAGGACAGGATAGTGCATGTTCTTTGTCTCTGAATTTTGTATATGTGC
TGTAATGTTGCTCTGAGGAAGCCCTGGAAAGTCTATCCCAACATATCCACATCTTATATTCAC
AAATTAAGCTGTAGTATGTACCCTAAGACGCTGCTAATTGACTGCCACTTCGCAACTCAGGGGCG
GCTGCATTTTAGTAATGGGTCAAATGATTACCTTTTATGATGCTTCCAAAGGTGCCTTGGCTTC
TCTTCCCAACTGACAAATGCCAAAGTTGAGAAAAATGATCATAATTTTAGCATAAACAGAGCAGT
CGGGGACACCGATTTTATAAATAAACTGAGCACCTTCTTTTAAACAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 6o

MASLGQILFWSIIISIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDIKLS
DIVIQWLKEGVLGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASLRLKNVQLTDAGTYKC
YIITSKKGKGNANLEYKTGAFSMPEVNVVDYNASSETLRCEAPRWFPQPTVVWASQVDQGANFSEVS
NTSFELNSENVTMKVVSVLYNVTINNTYSCMIENDIAKATGDIKVTSEIKRRSHLQLLNSKASL
CVSSFFAISWALLPLSPYLMLK

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 258-281

N-glycosylation sites.

amino acids 112-116, 160-164, 190-194, 196-200, 205-209, 216-220,
220-224

N-myristoylation sites.

amino acids 52-58, 126-132, 188-194

FIGURE 61

TGACGTCAGAATCACCAATGGCCAGCTATCCTTACCGGCAGGGCTGCCCAGGAGCTGCAGGACAAG
 CACCAGGAGCCCCCTCCGGGTAGCTACTACCCTGGACCCCCCAATAGTGGAGGGCAGTATGGTAGT
 GGGCTACCCCCTGGTGGTGGTTATGGGGGTCCCTGCCCTGGAGGGCCTTATGGACCACCAGCTGG
 TGGAGGGCCCTATGGACACCCCAATCCTGGGATGTTCCCCTCTGGAACCTCAGGAGGACCATATG
 GCGGTGCAGCTCCCGGGGGCCCCCTATGGTCAGCCACCTCCAAGTTCCTACGGTGCCAGCAGCCT
 GGGCTTTATGGACAGGGTGGCGCCCCCTCCCAATGTGGATCCTGAGGCCTACTCCTGGTTCCAGTC
 GGTGGACTCAGATCACAGTGGCTATATCTCCATGAAGGAGCTAAAGCAGGCCCTGGTCAACTGCA
 ATTGGTCTTCATTCAATGATGAGACCTGCCTCATGATGATAAACATGTTTGACAAGACCAAGTCA
 GGCCGCATCGATGCTACGGCTTCTCAGCCCTGTGGAAATTCATCCAGCAGTGGAAGAACCTCTT
 CCAGCAGTATGACCGGGACCGCTCGGGCTCCATTAGCTACACAGAGCTGCAGCAAGCTCTGTCCC
 AAATGGGCTACAACCTGAGCCCCAGTTACCCAGCTTCTGGTCTCCCGCTACTGCCACGCTCT
 GCCAATCCTGCCATGCAGCTTGACCGCTTCATCCAGGTGTGCACCCAGCTGCAGGTGCTGCAGA
 GGCCTTCCGGGAGAAGGACACAGCTGTACAAGCAACATCCGGCTCAGCTTCGAGGACTTCGTCA
 CCATGACAGCTTCTCGGATGCTATTGACCCAACCATCTGTGGAGAGTGGAGTGCACCAGGGACCTT
 TCCTGGCTTCTTAGAGTGAGAGAAGTATGTGGACATCTCTTCTTTCTGTCCCTCTAGAAGAAC
 ATTCTCCCTTGCTTGATGCAACACTGTTCCAAAAGAGGGTGGAGAGTCCATGCATCATAGCCACCA
 AATAGTGAGGACCGGGGCTGAGGCCACACAGATAGGGGCTGATGGAGGAGAGGATAGAAGTTGA
 ATGTCCTGATGGCCATGAGCAGTTGAGTGGCACAGCCTGGCACCAGGAGCAGGTCTTGTAAATGG
 AGTTAGTGTCCAGTCAGCTGAGCTCCACCTGATGCCAGTGGTGAGTGTTCATCGGCCTGTTACC
 GTTAGTACCTGTGTTCCCTCACCAGGCCATCCTGTCAAACGAGCCCATTTTCTCAAAGTGGAAT
 CTGACCAAGCATGAGAGAGATCTGTCTATGGGACCAGTGGCTTGGATTCTGCCACACCCATAAAT
 CCTTGTGTGTTAACTTCTAGCTGCCTGGGGCTGGCCCTGCTCAGACAAATCTGCTCCCTGGGCAT
 CTTTGGCCAGGCTTCTGCCCCCTGCAGCTGGGACCCCTCACTTGCTGCCATGCTCTGCTCGGCT
 TCAGTCTCCAGGAGACAGTGGTCACCTCTCCCTGCCAATACTTTTTTTAATTTGCATTTTTTTTC
 ATTTGGGGCCAAAAGTCCAGTGAAATTGTAAGCTTCAATAAAAGGATGAAACTCTGA

[illegible]

Important features of the protein:

amino acids 1-19

amino acids 147-150

amino acids 135-138, 150-153, 202-205, 271-274

amino acids 9-14, 15-20, 19-24, 33-38, 34-39, 39-44, 43-48, 61-66, 70-75, 78-83, 83-88, 87-92, 110-115

FIGURE 63

CAGGATGCAGGGCCGCGTGGCAGGGAGCTGCGCTCCTCTGGGCCTGCTCCTGGTCTGTCTTCATC
TCCCAGGCCTCTTTGCCCGGAGCATCGGTGTTGTGGAGGAGAAAGTTTCCCAAACTTCGGGACC
AACTTGCTCAGCTCGGACAACCTTCCTCCACTGGCCCCCTCTAACTCTGAACATCCGCAGCCCGC
TCTGGACCCTAGGTCTAATGACTTGGCAAGGGTTCTCTGAAGCTCAGCGTGCCTCCATCAGATG
GCTTCCCACCTGCAGGAGGTTCTGCAGTGCAGAGGTGGCCTCCATCGTGGGGGCTGCCTGCCATG
GATTCTGGCCCCCTGAGGATCCTTGGCAGATGATGGCTGCTGCGGCTGAGGACCGCTGGGGGA
AGCGCTGCCTGAAGAACTCTCTTACCTCTCCAGTGCTGCGGCCCTCGCTCCGGGCAGTGGCCCTT
TGCCTGGGGAGTCTTCTCCCGATGCCACAGGCCTCTCACCTGAGGCTTCACTCCTCCACCAGGAC
TCGGAGTCCAGACGACTGCCCCGTTCTAATTCAGTGGGAGCCGGGGGAAAAATCCTTTCCCAACG
CCCTCCCTGGTCTCTCATCCACAGGGTTCTGCCTGATCACCCCTGGGGTACCCTGAATCCCAGTG
TGTCTGGGGAGGTGGAGGCCCTGGGACTGGTTGGGGAACGAGGCCCATGCCACACCCTGAGGGA
ATCTGGGGTATCAATAATCAACCCCCAGGTACCAGCTGGGGAAATATTAATCGGTATCCAGGAGG
CAGCTGGGGAAATATTAATCGGTATCCAGGAGGCAGCTGGGGGAATATTAATCGGTATCCAGGAG
GCAGCTGGGGGAATATTCATCTATACCCAGGTATCAATAACCCATTTCTCCTGGAGTTCTCCGC
CCTCCTGGCTCTTCTTGGAACATCCCAGCTGGCTTCCCTAATCCTCCAAGCCCTAGGTTGCAGTG
GGGCTTAGAGCACGATAGAGGGAAACCAACATTGGGAGTTAGAGTCTGCTCCCGCCCCCTTGCTG
TGTGGGCTCAATCCAGGCCCTGTTAACATGTTTCCAGCACTATCCCCACTTTTCAGTGCCTCCCC
TGCTCATCTCCAATAAAATAAAAGCACTTATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAA

FIGURE 64

MQGRVAGSCAPLGILLVCLHLPGLFARSIGVVEEKVSQNFGTNLPQLGQPSSTGPSNSEHPQPAL
DPRSNDLARVPLKLSVPPSDGFPPAGGSQVQRWPPSWGLPAMDSWPPEDPWQMAAAAEDRLGEA
LPEELSYLSSAAALAPGSGPLPGESSPDATGLSPEASLLHQDSESRRRLPRSNSLGAGGKILSRP
PWSLIHRVLPDHPWGTNLNPSVSWGGGGPGTGWGTRPMPHPGEGIWGINNQPPGTSGWNINRYPGGS
WGNINRYPGGSWGNINRYPGGSWGNIHLYPGINNPFPPGVLRRPPGSSWNIPAGEFPNPPSPRLQWG

Important features of the protein:

Signal peptide:

amino acids 1-26

Casein kinase II phosphorylation sites.

amino acids 56-59, 155-158

N-myristoylation sites.

amino acids 48-53, 220-225, 221-226, 224-229, 247-252, 258-263,
259-264, 269-274, 270-275, 280-285, 281-286, 305-310

FIGURE 65

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCTCCATCCCAGGAGCGCAGTGGCCACTATGGGGTC
TGGGCTGCCCCCTTGTCTCCTCTTGACCCTCCTTGGCAGCTCACATGGAACAGGGCCGGGTATGA
CTTTGCAACTGAAGCTGAAGGAGTCTTTCTGACAAATTCCTCCTATGAGTCCAGCTTCCTGGAA
TTGCTTGAAAAGCTCTGCCTCCTCCTCCATCTCCCTTCAGGGACCAGCGTCACCCTCCACCATGC
AAGATCTCAACACCATGTTGTCTGCAACACATTGACAGCCATTGAAGCCTGTGTCTTCTTGGCCC
GGGCTTTTGGGCCGGGGATGCAGGAGGCAGGCCCCGACCCTGTCTTTCAGCAGGCCCCCACCTC
CTGAGTGGCAATAAATAAAATTCGGTATGCTG

FIGURE 66

MGSGPLVLVLLLTLLGSSHGTGPGMTLQLKLKESFLTNSSESSFLELLEKLCLLLHLPSTSVTL
HHARSQHVVNT

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 37-41

N-myristoylation sites.

amino acids 15-21, 19-25, 60-66

FIGURE 67

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCCGGGC
CAGGTGCCCCGTGCGCAGGTGCCCCCTGGCCGGAGATGCGGTAGGAGGGGCGAGCGCGAGAAGCCCC
TTCTCGGGCGCTGCCAACCCGCCACCCAGCCCCATGGCGAACCCCGGGCTGGGGCTGCTTCTGGCG
CTGGGCCTGCCGTTCCTGCTGGCCCGCTGGGGCCGAGCCTGGGGGCAAATACAGACCACTTCTGC
AAATGAGAATAGCACTGTTTTGCCTTCATCCACCAGCTCCAGCTCCGATGGCAACCTGCGTCCGG
AAGCCATCACTGCTATCATCGTGGTCTTCTCCCTCTTGGCTGCCTTGCTCCTGGCTGTGGGGCTG
GCACTGTTGGTGCGGAAGCTTCGGGAGAAGCGGCAGACGGAGGGCACCTACCGGCCCAGTAGCGA
GGAGCAGTTCTCCCATGCAGCCGAGGCCCGGGCCCCCTCAGGACTCCAAGGAGACGGTGCAGGGCT
GCCTGCCCATCTAGGTCCCCTCTCCTGCATCTGTCTCCCTTCATTGCTGTGTGACCTTGGGGAAA
GGCAGTGCCCTCTCTGGGCAGTCAGATCCACCCAGTGCTTAATAGCAGGGAAGAAGGTACTTCAA
AGACTCTGCCCCTGAGGTCAAGAGAGGATGGGGCTATTCACCTTTATATATTTATATAAAATTAG
TAGTGAGATGTAAAAAAAAAAAAAAAAAAAA

FIGURE 68

MANPGLGLLLALGLPFLLARWGRAWGQIQTTSANENSTVLPSTSSSSDGNLRPEAITAIIVVFS
LLAALLLAVGLALLVRKLREKRQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

Important features:**Signal peptide:**

amino acids 1-19

Transmembrane domain:

amino acids 56-80

N-glycosylation site.

amino acids 36-40

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 86-90

Tyrosine kinase phosphorylation site.

amino acids 86-94

N-myristoylation sites.

amino acids 7-13, 26-32

FIGURE 69

GCCAGGAATAACTAGAGAGGAACAATGGGGTTATTGAGAGGTTTTGTTTTCTCTTAGTTCGTGCCTGCTGCACCAG
 TCAAATACTTCCTTCATTAAGCTGAATAATAATGGCTTTGAAGATATTGTCATTGTTATAGATCCTAGTGTGCCAGAA
 GATGAAAAATAATTGAACAAATAGAGGATATGGTGACTACAGCTTCTACGTACCTGTTGAAGCCACAGAAAAAGA
 TTTTTTTTCAAAAAATGATCTATATTAATTCCTGAGAATTGGAAGGAAAAATCCTCAGTACAAAAGGCCAAAAATGAA
 AACCATAAACATGCTGATGTTATAGTTGCACCACCTACACTCCCAGGTAGAGATGAACCATACACCAAGCAGTTTACA
 GAATGTGGAGAGAAAGGCGAATACATTCACCTTCACCCCTGACCTTCTACTTGGAAAAAACAAAAATGAATATGGACCA
 CCAGGCCAACTGTTTGTCCATGAGTGGGCTCACCTCCGGTGGGGAGTGGTTGATGAGTACAATGAAGATCAGCCTTTC
 TACCGTGCTAAGTCAAAAAAATCGAAGCAACAAGGTGTTCCGCAGGTATCTCTGGTAGAAATAGAGTTTATAAGTGT
 CAAGGAGGCAGCTGTCTTAGTAGAGCATGCAGAATTGATTCTACACAAAACGTATGGAAAAAGATTGTCAATTCTTT
 CCTGATAAAGTACAAACAGAAAAAGCATCCATAATGTTTATGCAAAGTATTGATTCTGTTGTTGAATTTGTAAACGAA
 AAAACCCATAATCAAGAAGCTCCAAGCCTACAAACATAAAGTGCAATTTAGAAGTACATGGGAGGTGATTAGCAAT
 TCTGAGGATTTTAAAAACACCATACCCATGGTGACACCACCTCCTCCACCTGTCTTCTCATTGCTGAAGATCAGTCAA
 AGAATTTGTGCTTAGTTCCTTGATAAGTCTGGAAGCATGGGGGGTAAAGCACCGCTAAATCGAATGAATCAAGCAGCA
 AAACATTTCTGCTGCAGACTGTTGAAAAATGGATCCTGGGTGGGGATGGTTCACTTTGATAGTACTGCCACTATTGTA
 AATAAGCTAATCCAAATAAAAAAGCAGTGATGAAAGAAACACACTCATGGCAGGATTACCTACATATCCTCTGGGAGGA
 ACTTCCATCTGCTCTGGAATTAATATGCATTTAGGTGATTGGAGAGTACATTCCTCAACTCGATGGATCCGAAGTA
 CTGCTGCTGACTGATGGGAGGATAACACTGCAAGTTCCTGTATTGATGAAGTGAACAAAGTGGGGCCATTGTTCTAT
 TTTATTGCTTTGGGAAGAGCTGCTGATGAAGCAGTAATAGAGATGAGCAAGATAACAGGAGGAAGTCATTTTATGTT
 TCAGATGAAGCTCAGAACAATGGCCTCATTGATGCTTTTGGGGCTTACATCAGGAAATACTGATCTCTCCGAGAAG
 TCCCTTCAGCTCGAAAGTAAGGGATTAACTGAATAGTAATGCCCTGGATGAACGACACTGTCAATTTAGTAGTACA
 GTGGGAAAGGACACGTTCTTTCTCATCACATGGAACAGTCTGCCTCCAGTATTTCTCTCTGGGATCCAGTGGAAACA
 ATAATGGAAAAATTTACAGTGGATGCAACTTCCAAAATGGCCTATCTCAGTATTCAGGAACTGCAAAAGTGGGCACT
 TGGGCATACAATCTTCAAGCCAAAGCGAACCAGAAACATTAACATTTACAGTAACTTCTCGAGCAGCAAATTTCTCT
 GTGCCTCCAATCACAGTGAATGCTAAAATGAATAAGGACGTAACAGTTCCTCCAGCCCAATGATTGTTTACGAGAA
 ATTCTACAAGGATATGTACCTGTTCTTGGAGCCAAATGTGACTGCTTTTATTGAATCACAGAATGGACATACAGAAGTT
 TTGGAACTTTTGGATAATGGTGACGGCGCTGATTCTTTCAAGAATGATGGAGTCTACTCCAGGTATTTTACAGCATAI
 ACAGAAAATGGCAGATATAGCTTAAAGTTCGGGCTCATGGAGGAGCAACACTGCCAGGCTAAAATTACGGCTCCA
 CTGAATAGAGCCGCTACATACCAGGCTGGGTAGTGAACGGGGAAATGAAGCAAACCCGCAAGACCTGAAATTTGAT
 GAGGATACTCAGACCACCTTGGAGGATTTAGCCGAACAGCATCCGGAGGTGCATTTGTGGTATCACAACTCCCAAGC
 CTTCCTTGCCTGACCAATACCCACCAAGTCAATACAGACCTTGATGCCACAGTTCATGAGGATAAGATTATCTT
 ACATGGACAGCACCAGGAGATAATTTTATGTTGGAAAAAGTCAACGTTATATCATAAGAAATAGTGAAGTATTCTT
 GATCTAAGAGACAGTTTTGATGATGCTCTTCAAGTAAATACTACTGATCTGTCAACAAAGGAGGCCAACTCCAAAGGAA
 AGCTTTGCATTTAAACCAGAAAATATCTCAGAAGAAAATGCAACCCACATATTATTGCCATTAAAGTATAGATAAA
 AGCAATTTGACATCAAAAGTATCCAACATTGCACAAGTAACTTTGTTTATCCCTCAAGCAAATCCTGATGACATTGAT
 CCTACACCTACTCCTACTCTCTACTCTCTGATAAAAAGTCATAATTTCTGGAGTTAATATTTCTACGCTGGTATTG
 TCTGTGATTGGGTCTGTTGTAATTTGTTAACTTTATTTTAAAGTACCACCATTTGAACCTTAACGAAGAAAAAATCTTC
 AAGTAGACCTAGAAGAGAGTTTTTAAAAACAAAACAATGTAAGTAAAGGATATTCTGAATCTTAAATTCATCCCAT
 GTGTGATCATAACTCATAAAAATAATTTTAAAGATGTCGGAAGGAGTACTTTGATTAAATAAAAAACACTCATGGATA
 TGTAAAAAAGTGTCAAGATTAAAAATTAATAGTTTCATTTATTTGTTATTTTATTGTAAGAAATAGTGATGAACAAAG
 ATCCTTTTTTCACTGATACCTGGTTGTATATTATTTGATGCAACAGTTTTCTGAAATGATATTTCAAATTCATCAA
 GAAATTAATCATCTATCTGAGTAGTCAAAATACAAGTAAAGGAGAGCAAATAAACACATTTGGAAAAAATAAAAA
 AA

FIGURE 70

MGLFRGFVFLVLCLLHQSNSTFIKLNNGFEDIVIVIDPSVPEDEKIEQIEDMVTTASTYLFE
 ATEKRFFFKNVSILIPENWKENPQYKRPKHENHKHADVIVAPPTLPGRDEPYTKQFTECGEKGEY
 IHFTPDLLLGGKKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATRCISAGISGRN
 RVYKCQGGSCLSRACRIDSTTKLYGKDCQFFPDQVTEKASIMFQSIDSVVEFCNEKTHNQEAP
 SLQNIKCNFRSTWEVISNSEDFKNTIPMVTPPPPVFSLLKISQRIVCLVLDKSGSMGGKDRNLNR
 MNQAAKHFLQLTVENGSWGMVHFDSTATIVNKLIQIKSSDERNTLMAGLPTYPLGGTSICSGIK
 YAFQVIGELHSQLDGSEVLLLLTDGEDNTASSCIDEVKQSGAIVHFIALGRAADEAVIEMSKITGG
 SHFYVSDEAQNNGLIDAFGALTSGNTDLSQKSLQLESKGLTLNSNAWMNDTVIIDSTVGKDTFFL
 ITWNSLPPSISLWDPSTIMENFTVDATSKMAYLSIPGTAKVGTWAYNLQAKANPETLTITVTSR
 AANSSVPPITVNAKMNKDVNSFSPMIVYAEILQGYVPVLGANVTAFIESQNGHTEVLELLDNGA
 GADSFKNQGVYSRYFTAYTENGYSKVRHGGANTARLKLRPPLNRAAYIPGWVVNGEIEANPP
 RPEIDEDTQTTLDEFSRTASGGAFVVSQVPSLPLPDQYPPSQITDLDATVHEDKIIILTWAPGDN
 FDVGKVQRYIIRISASILDLRDSFDDALQVNTTDLSPKEANSKESFAFKPENISEENATHIFIAI
 KSIDKSNLTSKVSNIQVTLFIPQANPDDIDPTPTPTPTPTDKSHNSGVNISTLVLSVIGSVVI
 VNFILSTTI

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592,
 628-632, 811-815, 832-836, 837-841, 852-856, 896-900

FIGURE 71

CTCCTTAGGTGAAAACCCCTGGGAGTAGAGTACTGACAGCAAAGACCGGAAAGACCATACGTCCCCGGGACAGGGGTGA
CAACAGGTGTATCTTTTATCTCGTGTGTGGCTGCCTTCTATTTCAAGGAAAGACGCCAAGGTAATTTTGACCCA
GAGGAGCAATGATGTAGCCACCTCCTAACCTTCCCTTCTTGAACCCCAAGTTATGCCAGGATTTACTAGAGAGTGTCA
ACTCAACCAGCAAGCGGCTCCTTCGGCTTAACTTGTGGTTGGAGGAGAGAACCTTTGTGGGGCTGCGTTCTCTAGCA
GTGCTCAGAAGTGACTTGCTGAGGGTGGACCAGAAGAAAGGAAAGGTCCCCCTTGTCTGTGGCTGCACATCAGGAA
GGCTGTGATGGGAATGAAGGTGAAAACCTTGGAGATTTCACTTCAGTCATTGCTTCTGCCTGCAAGATCATCTTTAAA
AGTAGAGAAAGCTGCTCTGTGTGGTGGTTAACTCCAAGAGGCAGAACTCGTTCTAGAAGGAAATGGATGCAAGCAGCTC
CGGGGGCCCCAAACGCATGCTTCTGTGGTCTAGCCCAGGGAAGCCCTTCCGTGGGGGGCCCCGGCTTTGAGGGATGCC
ACCGGTTCTGGACGCATGGCTGATTCCTGAATGATGCTTCCCGGGGGCTGCTTGCCTGGATTCCCGGGTGGTG
GTTTTGCTGGTGCTCTGTGTGCTATCTGTCTGTACATGTTGGCTGCACCCCAAAGGTGACGAGGAGCAG
CTGGCACTGCCAGGGCCAAACAGCCACGGGGAAAGGGGTACCAGGCCGTCTTCAAGAGTGGGAGGAGCAGCAC
CGCAACTACGTGAGCAGCTGAAGCGGCAGATCGCACAGCTCAAGGAGGAGCTGCAGGAGAGGAGTGAGCAGCTCAGG
AATGGGCACTACCAAGCCAGCGATGCTGTGGCTGGGTCTGGACAGGAGCCCCCAGAGAAAACCCAGGCCGACCTC
CTGGCCTTCTGCACTCGCAGGTGGCAAGGCAGAGGTGAATGCTGGCGTCAAGCTGGCCACAGAGTATGCAGCAGTG
CCTTTGATAGCTTTACTCTACAGAAGGTGACCAGCTGGAGACTGGCCTTACCCGCCACCCGAGGAGAAGCTGTG
AGGAAGGACAAGCGGATGAGTTGGTGGAGCCATTGAATCAGCCTTGGAGACCTGAACAATCTGCAGAGAACAGC
CCCAATCACCGTCTTACAGCGCCTCTGATTTCTAGAGAGGATCTACCGAACAGAAAGGGACAAAGGGACATTGTAT
GAGCTCACCTTCAAAGGGGACCACAAACACGAATTCAAACCGCTCATCTTATTTCGACCAATCAGCCCCATCATGAAA
GTGAAAATGAAAAGCTCAACATGGCCAAACAGCTTATCAATGTTATCGTGCCTCTAGCAAAAAGGGTGGACAGTTTC
CGGCAGTTTATGCAGAAATTCAGGGAGATGTGATTGAGCAGGATGGGAGAGTCCATCTCACTGTTGTTTACTTTGGG
AAAGAAGAAATAAATGAAGTCAAAGGAATACTTGAAAACACTTCAAAGCTGCCAACTTCAGGAATTTTACCTTCATC
CAGCTGAATGGAGAAATTTCTCGGGGAAAGGGACTTGATGTTGGAGCCCGCTTCTGGAAGGGAAGCAACGTCTCTCTC
TTTTCTGTGATGTGGACATCTACTTCACATCTGAATTCCTCAATACGTGTAGGCTGAATACAGCCAGGGAAGAAG
GTATTTTATCCAGTTCTTTTCAGTCAGTACAATCCTGGCATAATATACGGCCACCATGATGCAGTCCCTCCCTTGGAA
CAGCAGCTGGTCATAAAGAAAGAACTGGATTTTGGAGAGACTTTGGATTGGGATGACGTGTGAGTATCGGTGAGCAG
TTTCAATATAGGTGGTTTGTCTGACATCAAAGCTGGGGCGGAGAGGATGTGCACCTTTATCGCAAGTATCTC
CACAGCAACCTCATAGTGGTACGGACGCCTGTGGGAGGACTCTTCCACCTCTGGCATGAGAAGCGCTGCATGGACGAG
CTGACCCCGAGCAGTACAAGATGTGCATGCAGTCCAAGGCCATGAACGAGGCATCCACGGCCAGCTGGGCATGCTG
GTGTTCAAGGCACGAGATAGAGGCTCACCTTCGCAACAGAAACAGAAAGCAAGTAGCAAAAAACATGAACCTCCAGA
GAAGGATTTGGGAGACACTTTTTCTTCTTTTGAATTAAGTGGCTGCAACAGAGAAAAAGCTTCCATAAA
GGACGACAAAAGAAATGGACTGATGGGTGAGAGATGAGAAAGCCCTCCGATTTCTCTGTGGGGCTTTTACAAACAGA
AATCAAAATCTCCGCTTTGCTGCAAAAGTAACCCAGTTGCACCTGTGAAGTGTCTGACAAAGGCAGAAATGCTTGTG
AGATTATAAGCCTAATGTTGTGGAGGTTTTGATGTTGTTACAATACACTGAGACCTGTTGTTTGTGTGCTCATTGA
AATATTCATGATTTAAGAGCAGTTTTGTAAAAAATTCATTAGCATGAAAGGCCAAGCATATTTCTCCTCATATGAATGA
GCCTATCAGCAGGGCTCTAGTTTCTAGGAATGCTAAAAATATCAGAAGGCAGGAGAGGATAGGCTTATATGATACT
AGTGAGTACATTAAGTAAATAAAATGGACCAGAAAAAGAAAGAAACCATAAATATCGTGTATATTTCCCAAGAT
TAACCAAAAAATATCTGCTTATCTTTTGGTTGCTCTTTAACTGTCTCCGTTTTTTCTTTTATTTAAAAATGCAC
TTTTTCCCTTGTGAGTTATAGTCTGCTTATTTAATTACCACTTTGCAAGCCTTACAAGAGAGCACAAGTTGCGCTAC
ATTTTTATATTTTTAAGAAGATACTTTGAGATGCATTATGAGAATTTTCAAGTCAAAATGATGCCATAT
CCAAGGACATGCCAAATGCTGATTCTGTGAGGCATGAATGTCAGGCATTGAGACATAGGGAAGGAATGGTTGTACT
AATACAGAGCTACAGATACTTTCTCTGAAGAGTATTTTGAAGAGGAGCAACTGAACACTGGAGGAAAAAGAAATGAC
ACTTTCTGCTTACAGAAAAGGAACTCATTGAGACTGGTGATATCGTGATGTACCTAAAAGTCAAGAAACCATTTTT
CTCCTCAGAAGTAGGGACCGCTTTCTACCTGTTTAAATAAACCAAGTATACCGTGTGAACCAACAACTCTTTTC
AAAACAGGGTCTCCTCCTGGCTTCTGGCTTCCATAAGAAGAAATGGAGAAAAATATATATATATATATATATTTGT
GAAAGATCAATCCATCTGCCAGAACTAGTGGGATGGAAGTTTTGCTACATGTTATCCACCCAGGCCAGGTGGAAG
TAACTGAATTTATTTTTAAATTAAGCAGTTCTACTCAATCACCAGATGCTTCTGAAAATTGCATTTTATACCATTT
CAAATATTTTTTAAAAATAAATACAGTTAATAGAGTGGTTTCTTCTCATGTGAAAATTTATAGCCAGCACCAG
ATGCATGAGCTAATTTCTTTGAGTCTTGTCTGTTTGTCTCACAGTAACTCATTTGTTTAAAGCTTCAAGAAC
ATTCAAGCTGTTGGTGTGTTAAAAATGCATTGTATTGATTGTACTGGTAGTTTATGAAATTTAATTAACACAGG
CCATGAATGGAAGGTGATTGACAGCTAATAAATATGATTTGTGGATATGAA

FIGURE 72

MMVRRGLLAWISRNVVLLVLLCCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQEW
EQHRNYVSSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGDRSPPEKTQADLLAFLHSQV
AEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPVKDKRDELVEAIESALETLN
ENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSPIMKVNEKLN
TLINIVPLAKRVDKFRQFMQNFREMCIEQDGRVHLTVVYFGKEEINEVKGILENTSKAANFR
TFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNTCRLNTQPGKKVFYPV
LFSQYNPGIYGHHDVAPPLEQQLVIKKETGFWRDFGFMTCQYRSDFINIGGFDDIKGWG
GEDVHLYRKYLHSNLIVVRTPVRGLFHLWHEKRCMDLTPEQYKCMQSKAMNEASHGQLG
MLVFRHEIEAHLRKQKQKTSSKKT

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 315-319, 324-328

N-myristoylation sites.

amino acids 96-102, 136-142, 212-218, 311-317, 339-345, 393-399

Amidation site.

amino acids 377-381

FIGURE 74

MLFSALLLEVIWILAADGGQHWTYEGPHGQDHWFPASYPECGNNAQSPIDIQTDSVTFDPLPALQ
PHGYDQPGTEPLDLHNNGHTVQLSLPSTLYLGGLPRKYVAAQLHLHWGQKSGPGGSEHQINSEAT
FAELHIVHYDSDSYDSLSEAAERPQGLAVLGILIEVGETKNIAYEHILSHLHEVRHKDQKTSVPP
FNLRELLPKQLGQYFRYNGSLTTPPCYQSVLWTVFYRRSQISMEQLEKLQGTLFSTEEEPSKLLV
QNYRALQPLNQRMVFASFIQAGSSYTTGEMLSLGVGILVGCLCLLLAVYFIARKIRKKRLENRKS
VVFTSAQATTEA

Important features of the protein:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 291-310

N-glycosylation site.

amino acids 213-216

Eukaryotic-type carbonic anhydrases proteins

amino acids 197-245, 104-140, 22-69

FIGURE 75

TGCCGCTGCCGCCGCTGCTGCTGTTGCTCCTGGCGGCGCCTTGGGGACGGGCAGTTCCCTGTGTC
 TCTGGTGGTTTGCCTAAACCTGCAAACATCACCTTCTTATCCATCAACATGAAGAATGTCCCTACA
 ATGGACTCCACCAGAGGGTCTTCAAGGAGTTAAAGTTACTTACACTGTGCAGTATTTTCATCACAA
 ATTGGCCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTTCTGTTGTCTGACAGCTCC
 AGAGAAGTGGAAGAGAAATCCAGAAGACCTTCCTGTTTCCATGCAACAAATATACTCCAATCTGA
 AGTATAACGTGTCTGTGTTGAATACTAAATCAAACAGAACGTGGTCCCAGTGTGTGACCAACCAC
 ACGCTGGTGTCTCACCTGGCTGGAGCCGAACACTCTTTACTGCGTACACGTGGAGTCCCTTCGTCCC
 AGGGCCCCCTCGCCGTGCTCAGCCTTCTGAGAAGCAGTGTGCCAGGACTTTGAAAGATCAATCAT
 CAGAGTTCAAGGCTAAAATCATCTTCTGGTATGTTTTGCCCATATCTATTACCGTGTTTCTTTTT
 TCTGTGATGGGCTATTCCATCTACCGATATATCCACGTTGGCAAAGAGAAACACCCAGCAAATTT
 GATTTTGATTTATGGAATGAATTTGACAAAAGATTCTTTGTGCCTGCTGAAAAAATCGTGATTA
 ACTTTATCACCTCAATATCTCGGATGATTCTAAAATTTCTCATCAGGATATGAGTTTACTGGGA
 AAAAGCAGTGATGTATCCAGCCTTAATGATCCTCAGCCCAGCGGGAACCTGAGGCCCCCTCAGGA
 GGAAGAGGAGGTGAAACATTTAGGGTATGCTTCGCATTTGATGGAAATTTTTTGTGACTCTGAAG
 AAAACACGGAAGGTACTTCTCTCACCCAGCAAGAGTCCCTCAGCAGAACAATACCCCGGATAAA
 ACAGTCATTGAATATGAATATGATGTCAGAACCACTGACATTTGTGCGGGGCCTGAAGAGCAGGA
 GCTCAGTTTGCAGGAGGAGGTGTCCACACAAGGAACATTATTGGAGTCGCAGGCAGCGTTGGCAG
 TCTTGGGCCCCGAAACGTTACAGTACTCATACCCCCTCAGCTCCAAGACTTAGACCCCCTGGCG
 CAGGAGCACACAGACTCGGAGGAGGGGCCGGAGGAAGAGCCATCGACGACCCTGGTCGACTGGGA
 TCCCCAACTGGCAGGCTGTGTATTCCCTTCGCTGTCCAGCTTCGACCAGGATTGAGAGGGCTGCG
 AGCCTTCTGAGGGGGATGGGCTCGGAGAGGAGGGTCTTCTATCTAGACTCTATGAGGAGCCGGCT
 CCAGACAGGCCACCAGGAGAAAATGAAACCTATCTCATGCAATTCATGGAGGAATGGGGTTATA
 TGTGCAGATGGAAAACTGATGGCCAACACTTCCTTTTGCCTTTTGTTCCTGTGCAAAACAAGTGAG
 TCACCCCTTTGATCCCAGCCATAAAGTACCTGGGATGAAAGAAGTTTTTCCAGTTTGTGAGTGT
 CTGTGAGAATTACTTATTTCTTTTCTATTCTCATAGCACGTGTGTGATTGGTTCATGCATGTA
 GGTCTCTTAACAATGATGGTGGGCCTCTGGAGTCCAGGGGCTGGCCGGTTGTTCTATGCAGAGAA
 AGCAGTCAATAAATGTTGCCAGACTGGGTGCAGAATTTATTCAGGTGGGTGT

FIGURE 76

MSYNGLHQRVFKELKLLTLCSSISQIGPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQQIY
 SNLKYNVSVLNTKSNRTWSQCVTNHTLVLTWLEPNTLYCVHVESFVPGPPRAQFSEKQCARTLK
 DQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILYGNEDKRFFVPAEK
 IVINFITLNISSDKISHQDMSLLGKSSDVSSLNDPQPSGNLRPPQEEEEVKHLGYASHLMEIFC
 DSEENTEGTSLTQQESLSRTIPPDKTVIEYEDVRTTDICAGPEEQELSLQEEVSTQGTLLESQA
 ALAVLGPOTLQYSYTPQLQDLPLAQEHTDSEEGPEEEPSTTLVDWDPQTGRLCIPSLSSFDQDS
 EGCEPSEGDGLGEEGLLSRLYEPPAPDRPPGENETYLMQFMEEWGLYVQMEN

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 140-163

N-glycosylation sites.

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

FIGURE 77

GAGGAGCGGGCCGAGGACTCCAGCGTGCCAGGTCTGGCATCCTGCACTTGCTGCCCTCTGACAC
 CTGGGAAGATGGCCGGCCCGTGGACCTTCACCCTTCTCTGTGGTTTGCTGGCAGCCACCTTGATC
 CAAGCCACCCTCAGTCCCAGTGCAGTTCTCATCCTCGGCCAAAAGTCATCAAAGAAAAGCTGAC
 ACAGGAGCTGAAGGACCACAACGCCACCAGCATCCTGCAGCAGCTGCCGCTGCTCAGTGCCATGC
 GGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGGCAGCCTGGTGAACACCGTCCTGAAGCACATC
 ATCTGGCTGAAGGTCATCACAGCTAACATCCTCCAGCTGCAGGTGAAGCCCTCGGCCAATGACCA
 GGAGCTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCAACACGCCCTGGTCAAGACCA
 TCGTGGAGTTCCACATGACGACTGAGGCCCAAGCCACCATCCGCATGGACACCAGTGCAAGTGGC
 CCCACCCGCTGGTCCTCAGTGACTGTGCCACCAGCCATGGGAGCCTGCGCATCCAATGCTGTA
 TAAGCTCTCCTTCTGGTGAACGCCTTAGCTAAGCAGGTGATGAACCTCCTAGTGCCATCCCTGC
 CCAATCTAGTGAAAAACCAGCTGTGTCCCGTGATCGAGGCTTCCTTCAATGGCATGTATGCAGAC
 CTCCTGCAGCTGGTGAAGGTGCCATTTCCTCAGCATTGACCGTCTGGAGTTTGACCTTCTGTA
 TCCTGCCATCAAGGGTGACACCATTAGCTCTACCTGGGGGCCAAGTTGTTGGACTCACAGGGAA
 AGGTGACCAAGTGGTTCAATAACTCTGCAGCTTCCTTGACAATGCCACCCTGGACAACATCCCG
 TTCAGCCTCATCGTGAGTCAGGACGTGGTGAAGCTGCAGTGGCTGCTGTGCTCTCTCCAGAAGA
 ATTCATGGTCCTGTTGGACTCTGTGCTTCCTGAGAGTGCCCATCGGCTGAAGTCAAGCATCGGGC
 TGATCAATGAAAAGGCTGCAGATAAGCTGGGATCTACCCAGATCGTGAAGATCCTAACTCAGGAC
 ACTCCCGAGTTTTTTATAGACCAAGGCCATGCCAAGGTGGCCCAACTGATCGTGCTGGAAGTGTT
 TCCCTCCAGTGAAGCCCTCCGCCCTTGTTCACCTGGGCATCGAAGCCAGCTCGGAAGCTCAGT
 TTTACACCAAAGGTGACCAACTTATACTCAACTTGAATAACATCAGCTCTGATCGGATCCAGCTG
 ATGAACTCTGGGATTGGCTGGTTCCAACCTGATGTTCTGAAAAACATCATCACTGAGATCATCCA
 CTCCATCCTGCTGCCGAACCAGAATGGCAAATTAAGATCTGGGGTCCCAGTGTCATTGGTGAAGG
 CCTTGGGATTCGAGGCAGCTGAGTCTCACTGACCAAGGATGCCCTTGTGCTTACTCCAGCCTCC
 TTGTGGAAACCCAGCTCTCCTGTCTCCAGTGAAGACTTGGATGGCAGCCATCAGGGAAGGCTGG
 GTCCCAGCTGGGAGTATGGGTGTGAGCTCTATAGACCATCCCTCTCTGCAATCAATAAACACTTG
 CCTGTGAAAAA

FIGURE 78

MAGPWTFLLCGLLAATLIQATLSPTAVLILGPKVIKEKLTQELKDNATSILQQLPLLSAMREK
PAGGIPVLGSLVNTVLKHIIWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPLVKTIVE
FHMTTEAQATIRMDTSASGPTRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMNLLVPSLPNL
VKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPAIKGDITQLYLGAKLLDSQGKVT
KWFNNSAASLTMPITLDNIPFSLIVSQDVVKAAVAVALSPPEEFMVLLDSVLPESAHLKSSIGLIN
EKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSSEALRPLFTLGIEASSEAQFYT
KGDQLILNLNLISSDRIQLMNSGIGWFQPDVLKNIITEIIHSILLPNQNGKLRSGVPVSLVKALG
FEAAESSLTkdalvltpaslwkpsfpvsq

Important features of the protein:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 48-51, 264-267, 401-404

Glycosaminoglycan attachment site.

amino acids 412-415

LBP / BPI / CETP family proteins.

amino acids 407-457

FIGURE 79

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGAGC
TTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTTCTACTGAGAGGTCTGCCATGGCCTCT
CTTGGCCTCCAACTTGTGGGCTACATCCTAGGCCTTCTGGGGCTTTTGGGCACACTGGTTGCCAT
GCTGCTCCCCAGCTGGAAAACAAGTTCTTATGTCGGTGCCAGCATTGTGACAGCAGTTGGCTTCT
CCAAGGGCCTCTGGATGGAATGTGCCACACACAGCACAGGCATCACCCAGTGTGACATCTATAGC
ACCCTTCTGGGCCTGCCCCTGACATCCAGGCTGCCAGGCCATGATGGTGACATCCAGTGCAAT
CTCTCCCTGGCCTGCATTATCTCTGTGGTGGGCATGAGATGCACAGTCTTCTGCCAGGAATCCC
GAGCCAAAGACAGAGTGGCGGTAGCAGGTGGAGTCTTTTTTCATCCTTGGAGGCCTCCTGGGATT
ATTCTGTGGCTGGAATCTTCATGGGATCTACGGGACTTCTACTCACCAGTGGTGCCTGACAG
CATGAAATTTGAGATTGGAGAGGCTCTTTACTTGGGCATTATTTCTTCCCTGTTCTCCCTGATAG
CTGGAATCATCCTCTGCTTTTCTGCTCATCCAGAGAAATCGCTCCAACTACTACGATGCCTAC
CAAGCCCAACCTCTTGCCACAAGGAGCTCTCCAAGGCCTGGTCAACCTCCCAAAGTCAAGAGTGA
GTTCAATTCCTACAGCCTGACAGGGTATGTGTGAAAGAACCAGGGGCCAGAGCTGGGGGTGGCTG
GGTCTGTGAAAAACAGTGGACAGCACCCCGAGGGCCACAGGTGAGGGACACTACCACTGGATCGT
GTCAGAAGGTGCTGCTGAGGATAGACTGACTTTGGCCATTGGATTGAGCAAAGGCAGAAATGGGG
GCTAGTGTAAACAGCATGCAGGTTGAATTGCCAAGGATGCTCGCCATGCCAGCCTTTCTGTTTTCC
TCACCTTGCTGCTCCCTGCCCTAAGTCCCCAACCCCTCAACTTGAAACCCATTCCCTTAAGCCA
GGACTCAGAGGATCCCTTTGCCCTCTGGTTTACCTGGGACTCCATCCCCAAACCCACTAATCACA
TCCCACTGACTGACCCTCTGTGATCAAAGACCCTCTCTCTGGCTGAGGTTGGCTCTTAGCTCATT
GCTGGGGATGGGAAGGAGAAGCAGTGGCTTTTGTGGGCATTGCTCTAACCTACTTCTCAAGCTTC
CCTCCAAAGAACTGATTGGCCCTGGAACCTCCATCCCACTCTTGTATGACTCCACAGTGTCCA
GACTAATTTGTGCATGAACTGAAATAAAACCATCCTACGGTATCCAGGGAACAGAAAGCAGGATG
CAGGATGGGAGGACAGGAAGGCAGCCTGGGACATTTAAAAAATA

FIGURE 8o

MASLGQLVGYILGLLGTLVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGITQCD
 IYSTLLGLPADIQAAQAMMTSSAIISSLACIISVVGMRCTVFCQESRAKDRVAVAGGVFFILGGL
 LGFIPVAWNHLHGILRDFYSPLVPDSMKFEIGEALYLGIISSLFSLIAGIILCFSCSSQRNRSNYY
 DAYQAQPLATRSSPRPGQPPKVKSEFNYSYSLTGYV

Important features of the protein:

Signal peptide:

amino acids 1-24

Transmembrane domains:

amino acids 82-102, 117-140, 163-182

N-glycosylation site.

amino acids 190-193

PMP-22 / EMP / MP20 family proteins.

amino acids 46-59

FIGURE 81

CCCACGCGTCCGCGCCTCTCCCTTCTGCTGGACCTTCCTTCGTCTCTCCATCTCTCCCTCCTTTC
 CCCGCGTTCTCTTTCCACCTTTCTCTTCTTCCACCTTAGACCTCCCTTCCTGCCCTCCTTTTCCT
 GCCCACCGCTGCTTCTTGGCCCTTCTCCGACCCCGCTCTAGCAGCAGACCTCCTGGGGTCTGTGG
 GTTGATCTGTGGCCCTGTGCCTCCGTGTCCTTTTCGTCTCCCTTCCTCCCGACTCCGCTCCCGG
 ACCAGCGGCTGACCTGGGGAAGGATGTTCCCGAGGTGAGGGTCTCTCCTCCTTGCTGGGA
 CTCGCGCTGCTCTGGTTCCCCCTGGACTCCACGCTCGAGCCCGCCAGACATGTTCTGCCTTTT
 CCATGGGAAGAGATACTCCCCCGGCGAGAGCTGGCACCCCTACTTGGAGCCACAAGGCCTGATGT
 ACTGCCTGCGCTGTACCTGCTCAGAGGGCGCCCATGTGAGTTGTTACCGCCTCCACTGTCCGCCT
 GTCCACTGCCCCAGCCTGTGACGGAGCCACAGCAATGCTGTCCCAAGTGTGTGGAACCTCACAC
 TCCCTCTGGACTCCGGGCCCCACCAAAGTCTGCCAGCACAACGGGACCATGTACCAACACGGAG
 AGATCTTCAGTGCCCATGAGCTGTTCCTTCCCGCTGCCCAACCAGTGTGTCTCTGCAGCTGC
 ACAGAGGGCCAGATCTACTGCGGCCCTACAACCTGCCCCGAACCAGGCTGCCAGCACCCCTCCC
 ACTGCCAGACTCCTGCTGCCAAGCCTGCAAAGATGAGGCAAGTGAAGCAATCGGATGAAGAGGACA
 GTGTGCAGTCGCTCCATGGGGTGAGACATCCTCAGGATCCATGTTCCAGTGATGCTGGGAGAAA
 AGAGGGCCGGGCACCCAGCCCCCACTGGCCTCAGCGCCCTCTGAGCTTCATCCCTCGCCACTT
 CAGACCCAAGGGAGCAGGCAGCACAACTGTCAAGATCGTCTGAAGGAGAAACATAAGAAAGCCT
 GTGTGCATGGCGGGAAGACGTACTCCACGGGGAGGTGTGGCACCCGGCCTTCCGTGCCTTCGGC
 CCCTTGCCCTGCATCCTATGCACCTGTGAGGATGGCCGCCAGGACTGCCAGCGTGTGACCTGTCC
 CACCGAGTACCCCTGCCGTACCCCGAGAAAGTGGCTGGGAAGTGTGCAAGATTTGCCAGAGG
 ACAAGCAGACCTGGCCACAGTGAGATCAGTTCTACCAGGTGTCCCAAGGCACCGGGCCGGGTC
 CTCGTCCACACATCGGTATCCCCAAGCCCAGACAACCTGCGTCGCTTTGCCCTGGAACACGAGGC
 CTCGGACTTGGTGGAGATCTACCTCTGGAAGCTGGTAAAAGATGAGGAACTGAGGCTCAGAGAG
 GTGAAGTACCTGGCCCAAGGCCACACAGCCAGAATCTTCCACTTGACTCAGATCAAGAAAGTCAG
 GAAGCAAGACTTCCAGAAAGAGGCACAGCACTTCCGACTGCTCGCTGGCCCCACGAAGGTCAC
 GGAACGTCTTCTAGCCAGACCTGGAGCTGAAGGTCACGGCCAGTCCAGACAAAGTGACCAAG
 ACATAACAAAGACCTTAACAGTTGCAGATATGAGCTGTATAATTGTTGTTATTATATATTAATAAA
 TAAGAAGTTGCATTACCCTCAAAAAAAAAAAAAAAAAAAAAA

FIGURE 82

MVPEVRVLSSLLGLALLWFPLDSHARARPDMFCLFHGKRYSPGESWHPYLEPQGLMYCLRTCSE
 GAHVSCYRLHCPPVHCPQVTEPQQCCPKVEPHTPSGLRAPPKSCQHNGTMYQHGEIFSAHELF
 PSRLPNQCVLCSCTEGQIYCGLTTCPEPGCPAPLPLPDSCCQACKDEASEQSDEEDSVQSLHGVR
 HPQDPCSSDAGRKRGPPTAPTGLSAPLSFIPRHFRPKGAGSTTVKIVLKEKHKACVHGGKTYS
 HGEVWHPAFRAFGPLPCILCTCEDGRQDCQRVTCPTTEPCRHPKVKAGKCKICPEDKADPGHSE
 ISSTRCPKAPGRVLVHTSVSPSPDNLRRFALEHEASDLVEIYLWKLVKDEETEAQRGEVPGPRPH
 SQNLPLSDQESQEARLPERGTALPTARWPFRSLERLPSDPGAEGHGQSRQSDQDITKT

Signal peptide:

amino acids 1-25

FIGURE 83

GACAGCTGTGTCTCGATGGAGTAGACTCTCAGAACAGCGCAGTTTGCCCTCCGCTCACGCAGAGCCTCTCC
GTGGCTTCCGCACCTTGAGCATTAGGCCAGTTCTCCTCTTCTCTAATCCATCCGTCACCTCTCCTGTCA
TCCGTTTCCATGCCGTGAGGTCCATTACAGAACACATCCATGGCTCTCATGCTCAGTTTGGTTCTGAGTC
TCCTCAAGCTGGGATCAGGGCAGTGGCAGGTGTTTGGGCCAGACAAGCCTGTCCAGGCCTTGGTGGGGGAG
GACGCAGCATTCTCCTGTTTCTGTCTCCTAAGACCAATGCAGAGGCCATGGAAGTCCGGTTCTTCAGGGG
CCAGTTCTCTAGCGTGGTCCACCTCTACAGGGACGGGAAGGACCAGCCATTTATGCAGATGCCACAGTATC
AAGGCAGGACAAAACCTGGTGAAGGATTCTATTGCGGAGGGGCGCATCTCTCTGAGGCTGGAACATTACT
GTGTTGGATGCTGGCCTCTATGGGTGCAGGATTAGTTCCAGTCTTACTACCAGAAGGCCATCTGGGAGCT
ACAGGTGTGAGCACTGGGCTCAGTTCCCTCTCATTTCATCAGGGATATGTTGATAGAGACATCCAGCTAC
TCTGTGAGTCCCTCGGGCTGGTTCCCCCGGCCACAGCGAAGTGGAAGGTCCACAAGGACAGGATTGTGCC
ACAGACTCCAGGACAAACAGAGACATGCATGGCCTGTTTGATGTGGAGATCTCTCTGACCGTCCAAGAGAA
CGCCGGGAGCATATCCTGTTCCATGCGGCATGCTCATCTGAGCCGAGAGGTGGAATCCAGGGTACAGATAG
GAGATACCTTTTTTCGAGCCTATATCGTGGCACCTGGCTACCAAAGTACTGGGAATACTCTGCTGTGGCTTA
TTTTTTGGCATTGTGGAAGTGAAGATTTCTTTCTCCAAATCCAGTGGAAATCCAGGCGGAAGTGGAGTGG
GAGAAGAAAGCACGGACAGGCAGAAATTGAGAGACCCCGGAAACACGCAGTGGAGGTGACTCTGGATCCAG
AGACGGCTCACCCGAAGCTCTGCGTTTCTGATCTGAAAACGTAAACCCATAGAAAAGTCCCCAGGAGGTG
CCTCACTCTGAGAAGAGATTTACAAGGAAGAGTGTGGTGGCTTCTCAGAGTTTCCAAGCAGGGAAACATTA
CTGGGAGGTGGACGGAGGACACAATAAAAGGTGGCGGTGGGAGTGTGCCGGATGATGTGGACAGGAGGA
AGGAGTACGTGACTTTGTCTCCCGATCATGGGTACTGGGTCTCAGACTGAATGGAGAACATTTGTATTTT
ACATTAAATCCCCGTTTTATCAGCGTCTTCCCCAGGACCCACCTACAAAAATAGGGGTCTTCTGAGCTA
TGAGTGTGGGACCATCTCCTTCTTCAACATAAATGACCAGTCCCTTATTTATACCCTGACATGTCGGTTTG
AAGGCTTATTGAGGCCCTACATTGAGTATCCGTCCTATAATGAGCAAAATGGAACTCCCATAGTCATCTGC
CCAGTCACCCAGGAATCAGAGAAAGAGGCCCTCTGGCAAAGGCCCTCTGCAATCCAGAGACAAGCAACAG
TGAGTCTCTCTCACAGGCAACACGCCCTTCCTCCCCAGGGGTGAAATGTAGGATGAATCACATCCCACAT
TCTTCTTTAGGGATATTAAGGTCTCTCTCCAGATCCAAAGTCCCGCAGCAGCCGCCAAGGTGGCTTCCA
GATGAAGGGGGACTGGCCTGTCCACATGGGAGTCAGGTGTCTGAGCTGCCCTGAGCTGGGAGGGAAGAAGG
CTGACATTACATTTAGTTTGCTCTCACTCCATCTGGCTAAGTGATCTTGAAATACCACCTCTCAGGTGAAG
AACCGTCAGGAATTCATCTCACAGGCTGTGGGTAGATTAAAGTAGACAAGGAATGTGAATAATGCTTAG
ATCTTATTGATGACAGAGTGTATCCTAATGGTTGTTCATTATATTACACTTTCAGTAAAAAAA

FIGURE 84

MALMLSLVLSLLKLGSGQWQVFGPDKPVQALVGEDAAFSCFLSPKTNAEAMEVRFFRGQFSSVVH
 LYRDGKDQPFMQMPQYQGRTKLVKDSIAEGRISLRLENITVLDAGLYGCRISQSYQKAIWELQ
 VSALGSVPLISITGYVDRDIQLLCQSSGWFPRTAKWKGPQGDLSTDSRTNRDMHGLFDVEISL
 TVQENAGSISCSMRHAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLFFGIVGLKIFFSK
 FQWKIQAELDWRRKHGQAEIRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAQEVPHSEKRF
 TRKSVVASQSFQAGKHYWEVDGGHNRWRVGVCRDDVDRRKEYVTLSPDHGYWVLRNLNGEHLFT
 LNPRFISVFPRTPTKIGVFLDYECGTISFFNINDQSLIYTLTCRFEGLLRPYIEYPSYNEQNGT
 PIVICPVTQESEKEASWQRASAIPESTNSSESSQATTPFLPRGEM

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 239-255

FIGURE 85

AACAGACGTTCCCTCGCGGCCCTGGCACCTCTAACCCAGACATGCTGCTGCTGCTGCTGCCCCT
GCTCTGGGGGAGGGAGAGGGCGGAAGGACAGACAAGTAAACTGCTGACGATGCAGAGTCCGTGA
CGGTGCAGGAAGGCCTGTGTGTCATGTGCCCTGCTCCTTCTCCTACCCCTCGCATGGCTGGATT
TACCTTGGCCCAGTAGTTCATGGCTACTGGTTCCGGAAGGGGCCAATACAGACCAGGATGCTCC
AGTGGCCACAACAACCCAGCTCGGGCAGTGTGGGAGGAGACTCGGGACCGATTCCACCTCCTTG
GGGACCCACATAACCAAGAATTGCACCCTGAGCATCAGAGATGCCAGAAGAAGTGATGCGGGGAGA
TACTTCTTTTCGTATGGAGAAAGGAAGTATAAAATGGAATTATAAACATCACCGGCTCTCTGTGAA
TGTGACAGCCTTGACCCACAGGCCCAACATCCTCATCCCAGGCACCCTGGAGTCCGGCTGCCCCC
AGAACTGACCTGCTCTGTGCCCTGGGCCTGTGAGCAGGGGACACCCCTATGATCTCCTGGATA
GGGACCTCCGTGTCCCCCTGGACCCCTCCACCACCCGCTCCTCGGTGCTCACCCCTCATCCCACA
GCCCCAGGACCATGGCACCAGCCTCACCTGTCAGGTGACCTTCCCTGGGGCCAGCGTGACCACGA
ACAAGACCGTCCATCTCAACGTGTCTACCCGCTCAGAACTTGACCATGACTGTCTTCCAAGGA
GACGGCACAGTATCCACAGTCTTGGGAAATGGCTCATCTCTGTCACTCCCAGAGGGCCAGTCTCT
GCGCCTGGTCTGTGCAGTTGATGCAGTTGACAGCAATCCCCCTGCCAGGCTGAGCCTGAGCTGGA
GAGGCCTGACCTGTGCCCCCTCACAGCCCTCAAACCCGGGGTGCTGGAGCTGCCTTGGGTGCAC
CTGAGGGATGCAGCTGAATTACCTGCAGAGCTCAGAACCTCTCGGCTCTCAGCAGGTCTACCT
GAACGTCTCCCTGCAGAGCAAAGCCACATCAGGAGTGACTCAGGGGGTGGTCGGGGGAGCTGGAG
CCACAGCCCTGGTCTTCTGTCTTCTGCGTCATCTTCGTTGTAGTGAGGTCTGCAGGAAGAAA
TCGGCAAGGCCAGCAGCGGGCGTGGGAGATACGGGCATAGAGGATGCAAACGCTGTCAGGGGTTC
AGCCTCTCAGGGGCCCTGACTGAACCTTGGGCAGAAGACAGTCCCCCAGACCAGCCTCCCCAG
CTTCTGCCCCTCCTCAGTGGGGGAAGGAGAGCTCCAGTATGCATCCCTCAGCTTCCAGATGGTG
AAGCCTTGGGACTCGCGGGGACAGGAGGCCACTGACACCGAGTACTCGGAGATCAAGATCCACAG
ATGAGAAACTGCAGAGACTCACCTGATTGAGGGATCACAGCCCTCCAGGCAAGGGAGAAGTCA
GAGGCTGATTCTTGTAGAATTAACAGCCCTCAACGTGATGAGCTATGATAACACTATGAATTATG
TGCAGAGTGAAAAGCACACAGGCTTTAGAGTCAAAGTATCTCAAACCTGAATCCACACTGTGCCC
TCCCTTTATTTTTTAACTAAAAGACAGACAAATTCCTA

FIGURE 86

MLLLLLPLLWGRERAEGQTSKLLTMQSSVTVQEGLCVHVPCSFSYPSHGWIIYPGPVVHGYWFREG
ANTDQDAPVATNNPARAVWEETRDRFHLLGDPHTKNCTLSIRDARRSDAGRYFFRMEKGSIKWNY
KHHRLSVNVTALTHRPNILIPGTLESGCPQNLTCSPWACEQGTPPMISWIGTSVSPLDPSTTRS
SVLTLLIPQPQDHGTS LTCQVTFPGASVTNKTVHLNVSYPPQNLMTVFQGDGTVSTVLGNGSSL
SLPEGQSLRLVCAVDAVDSNPPARLSLSWRGLTLCPSQPSNPGVLELPWVHLRDAAEFTCRAQNP
LGSQQVYLVNLSLQSKATSGVTQGVVGGAGATALVFLSFCVIFVVVRSCRKKSARPAAGVGD TGIE
DANAVRG SASQGPLTEPWAEDSPPDQPPPASARSSVGE GELQYASLSFQMVKPWDSRGQEATDTE
YSEIKIHR

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 351-370

FIGURE 87

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGGAGTGAAGGAGCTCTCTGTAC
 CCAAGGAAAGTGCAGCTGAGACTCAGACAAGATTACAATGAACCAACTCAGCTTCCTGCTGTTTC
 TCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTTCAAGGAATGGACCTGT
 TCTTCGTCTCCATCTCTGCCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCCTAGTGCATTTGA
 TGGCCTGTATTTTCTCCGCACTGAGAATGGTGTATCTACCAGACCTTCTGTGACATGACCTCTG
 GGGGTGGCGGCTGGACCCTGGTGGCCAGCGTGCATGAGAATGACATGCGTGGGAAGTGCACGGTG
 GCGGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCCAGAGGGGGACGGCAACTGGGC
 CAACTACAACACCTTTGGATCTGCAGAGGCGGCCACGAGCGATGACTACAAGAACCCTGGCTACT
 ACGACATCCAGGCCAAGGACCTGGGCATCTGGCACGTGCCCAATAAGTCCCCCATGCAGCACTGG
 AGAAACAGCTCCCTGCTGAGGTACCGCACGGACACTGGCTTCCTCCAGACACTGGGACATAATCT
 GTTTGGCATCTACCAGAAATATCCAGTGAAATATGGAGAAGGAAAGTGTGGACTGACAACGGCC
 CGGTGATCCCTGTGGTCTATGATTTTGGCGACGCCCAGAAAACAGCATCTTATTACTCACCTAT
 GGCCAGCGGGAATTCAGTGCGGGATTTGTTTCAGTTCAGGGTATTTAATAACGAGAGAGCAGCCAA
 CGCCTTGTGTGCTGGAATGAGGGTCACCGGATGTAACACTGAGCATCACTGCATTGGTGGAGGAG
 GATACTTTCCAGAGGCCAGTCCCCAGCAGTGTGGAGATTTTCTGGTTTTGATTGGAGTGGATAT
 GGAATCATGTTGGTTACAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCG
 TTGAGAGTTTTGTGGGAGGGAACCCAGACCTCTCCTCCCAACCATGAGATCCCAAGGATGGAGAA
 CAACTTACCCAGTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAAATCATATTGACTCAAGA
 AAAAAA

FIGURE 88

MNQLSFLLFLIATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTENGVI
YQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDRWSSQQGSKADYPEGDGNWANYNTFGSAEAAT
SDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSSLLRYRTDTGFLQTLGHNLFGIYQKYPVKYG
EGKCWTDNGPVI PVVYDFGDAQKTASYYSYGQREFTAGFVQFRVFNNERAANALCAGMRVTGCN
TEHHCIGGGGYFPEASPQCGDFSGFDWSGYGTHVGYSRSSREITEAAVLLFYR

Important features:**Signal peptide:**

amino acids 1-16

N-glycosylation site.

amino acids 163-167

Glycosaminoglycan attachment sites.

amino acids 74-78, 289-293

N-myristoylation sites.

amino acids 76-82, 115-121, 124-130, 253-259, 292-298

FIGURE 89

CTAGATTTGTCGGCTTGCGGGGAGACTTCAGGAGTCGCTGTCTCTGAACTTCCAGCCTCAGAGAC
CGCCGCCCTTGTCCCGAGGGCCATGGCCGGGTCTCAGGGCTTGTGCCCTCTCGCTTCCTGACG
CTCCTGGCGCATCTGGTGGTCGTATCACCTTATTCTGGTCCCGGGACAGCAACATACAGGCCTG
CCTGCCTCTCACGTTACCCCCGAGGAGTATGACAAGCAGGACATTCAGCTGGTGGCCGCGCTCT
CTGTCACCCTGGGCCTCTTTGCAGTGGAGCTGGCCGGTTTCCTCTCAGGAGTCTCCATGTTCAAC
AGCACCCAGAGCCTCATCTCCATTGGGGCTCACTGTAGTGCATCCGTGGCCCTGTCTTCTTCAT
ATTTCGAGCGTTGGGAGTGCACTACGTATTGGTACATTTTGTCTTCTGCAGTGCCCTTCCAGCTG
TCACTGAAATGGCTTTATTCGTCACCGTCTTTGGGCTGAAAAAGAAACCCTTCTGATTACCTTCA
TGACGGGAACCTAAGGACGAAGCCTACAGGGGCAAGGGCCGCTTCGTATTCTTGGAAGAAGGAAG
GCATAGGCTTCGGTTTTCCCTCGGAAACTGCTTCTGCTGGAGGATATGTGTTGGAATAAATACG
TCTTGAGTCTGGGATTATCCGCATTGTATTTAGTGCTTTGTAATAAAATATGTTTTGTAGTAACA
TTAAGACTTATATACAGTTTTAGGGGACAATTAAAAAAAAAAAA

FIGURE 90

MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACLPLTFTPEEYDKQDIQLVAALSVTLGLFA
VELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCSALPAVTEMALEFV
TVFGLKKKPF

Transmembrane domain:

amino acids 12-28 (type II), 51-66, 107-124

FIGURE 91

CTGGGACCCCGAAAAGAGAAGGGGAGAGCGAGGGGACGAGAGCGGAGGAGGAAGATGCAACTGAC
TCGCTGCTGCTTCGTGTTCTTGGTGCAGGGTAGCCTCTATCTGGTCATCTGTGGCCAGGATGATG
GTCCCTCCCGGCTCAGAGGACCCTGAGCGTGATGACCACGAGGGCCAGCCCCGGCCCCGGGTGCCT
CGGAAGCGGGGCCACATCTCACCTAAGTCCCGCCCCATGGCCAATTCCACTCTCCTAGGGCTGCT
GGCCCCGCCTGGGGAGGCTTGGGGCATTCTTGGGCAGCCCCCAACCGCCCGAACCACAGCCCCC
CACCCCTCAGCCAAGGTGAAGAAAATCTTTGGCTGGGGCGACTTCTACTCCAACATCAAGACGGTG
GCCCTGAACCTGCTCGTCCACAGGGAAGATTGTGGACCATGGCAATGGGACCTTCAGCGTCCACTT
CCAACACAATGCCACAGGCCAGGGAAACATCTCCATCAGCCTCGTGCCCCCAGTAAAGCTGTAG
AGTTCCACCAGGAACAGCAGATCTTCATCGAAGCCAAGGCCTCCAAAATCTTCAACTGCCGGATG
GAGTGGGAGAAGGTAGAACGGGGCCGCCGACCTCGCTTTGCACCCACGACCCAGCCAAGATCTG
CTCCCAGAGACCAGCTCAGAGCTCAGCCACCTGGAGCTGCTCCAGCCCTTCAAAGTCGTCTGTG
TCTACATCGCCTTCTACAGCACGGACTATCGGCTGGTCCAGAAGGTGTGCCAGATTACAACCTAC
CATAGTGATACCCCTACTACCCATCTGGGTGACCCGGGGCAGGCCACAGAGGCCAGGCCAGGGC
TGGAAGGACAGGCCTGCCATGCAGGAGACCATCTGGACACCGGGCAGGGAAGGGGTTGGGCCTC
AGGCAGGGAGGGGGGTGGAGACGAGGAGATGCCAAGTGGGGCCAGGGCCAAGTCTCAAGTGGCAG
AGAAAGGGTCCCAAGTGCTGGTCCCAACCTGAAGCTGTGGAGTGACTAGATCACAGGAGCACTGG
AGGAGGAGTGGGCTCTCTGTGCAGCCTCACAGGGCTTTGCCACGGAGCCACAGAGAGATGCTGGG
TCCCCGAGGCCTGTGGGCAGGCCGATCAGTGTGGCCCCAGATCAAGTCATGGGAGGAAGCTAAGC
CCTTGGTTCTTGCCATCCTGAGGAAAGATAGCAACAGGGAGGGGGAGATTTTCATCAGTGTGGACA
GCCTGTCAACTTAGGATGGATGGCTGAGAGGGCTTCTTAGGAGCCAGTCAGCAGGGTGGGGTGGG
GCCAGAGGAGCTCTCCAGCCCTGCCTAGTGGGCGCCCTGAGCCCCTTGTCTGTGTGCTGAGCATGG
CATGAGGCTGAAGTGGCAACCCTGGGGTCTTTGATGTCTTGACAGATTGACCATCTGTCTCCAGC
CAGGCCACCCCTTTCCAAAATTCCTCTTCTGCCAGTACTCCCCCTGTACCACCCATTGCTGATG
GCACACCCATCCTTAAGCTAAGACAGGACGATTGTGGTCTCCACACTAAGGCCACAGCCCATC
CGCGTGCTGTGTGTCCTCTTCCACCCCAACCCCTGCTGGCTCCTCTGGGAGCATCCATGTCCCCG
GAGAGGGGTCCCTCAACAGTCAGCCTCACCTGTGACACGGGGTTCTCCCGGATCTGGATGGCGC
CGCCCTCTCAGCAGCGGGCACGGGTGGGGCGGGGCCGGCCGAGAGCATGTGCTGGATCTGTTC
TGTGTGTCTGTCTGTGGGTGGGGGGAGGGGAGGGAAGTCTTGTGAAACCGCTGATTGCTGACTTT
TGTGTGAAGAATCGTGTTCTTGGAGCAGGAAATAAAGCTTGCCCCGGGGCA

FIGURE 92

MQLTRCCFVFLVQGSLYLVICGQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPKSRPMANSTL
LGLLAPPGEAWGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVALNLLVTGKIVDHGNGTF
SVHFQHNATGQGNISISLVPPSKAVEFHQEQQIFIEAKASKIFNCRMEWEKVERGRRTSLCTHDP
AKICSRDHAQSSATWSCSQPFKVVVCVYIAFYSTDYRLVQKVC PDYNYHS DTPYYPSG

Important features of the protein:

Signal peptide:

amino acids 1-14

N-glycosylation sites.

amino acids 62-65, 127-130, 137-140, 143-146

2-oxo acid dehydrogenases acyltransferase

amino acids 61-71

FIGURE 94

MTAAVFFGCAFIAGFPALALYVFTIAIEPLRIIFLIAGAFFWLVSLLISSLVWFMARVIIDNKDG
PTQKYLLIFGAFVSVYIQEMFREYAYKLLKKASEGLKSINPGETAPSMRLLAYVSGLGFGIMSGV
FSFVNTLSDSLGPGTVGIHGDSPOFFLYSAFMTLVIILLHVFWGIVFFDGCEKKKWGILLIVLLT
HLLVSAQTFISSYYGINLASAFIILVLMGTWAFLAAGGSCRSCLKLCLLCQDKNFFLYNQSR

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domains:

amino acids 32-51, 119-138, 152-169, 216-235

Glycosaminoglycan attachment site.

amino acids 120-123

Sodium:neurotransmitter symporter family protein

amino acids 31-65

FIGURE 95

AATTTTTCACCAGAGTAACTTGAGAAACCAACTGGACCTTGAGTATTGTACATTTTGCCTCGTG
GACCCAAAGGTAGCAATCTGAAACATGAGGAGTACGATTCTACTGTTTTGTCTTCTAGGATCAAC
TCGGTCATTACCACAGCTCAAACCTGCTTTGGGACTCCCTCCCACAAAAGTGGCTCCGGATCAGG
GAACACTACCAAACCAACAGCAGTCAAATCAGGTCTTTCCCTTCTTTAAGTCTGATACCATTAAACA
CAGATGCTCACACTGGGGCCAGATCTGCATCTGTAAATCCTGCTGCAGGAATGACACCTGGTAC
CCAGACCCACCCATTGACCCTGGGAGGGTTGAATGTACAACAGCAACTGCACCCACATGTGTTAC
CAATTTTGTACACAACCTGGAGCCCAGGGCACTATCCTAAGCTCAGAGGAATTGCCACAAATC
TTCAGGAGCCTCATCATCCATTCCCTTGTTCCCGGGAGGCATCCTGCCCACCAGTCAGGCAGGGGC
TAATCCAGATGTCCAGGATGGAAGCCTTCCAGCAGGAGGAGCAGGTGTAAATCCTGCCACCCAGG
GAACCCCAGCAGGCCGCCTCCCAACTCCCAGTGGCACAGATGACGACTTTGCAGTGACCACCCCT
GCAGGCATCCAAAGGAGCACACATGCCATCGAGGAAGCCACCACAGAATCAGCAAATGGAATTCA
GTAAGCTGTTTCAAATTTTTTCAACTAAGCTGCCTCGAATTTGGTGATACATGTGAATCTTATC
ATTGATTATATTATGGAATAGATTGAGACACATTGGATAGTCTTAGAAGAAATTAATTCCTAATT
TACCTGAAAATATTCTTGAAATTCAGAAAATATGTTCTATGTAGAGAATCCCAACTTTTAAAAA
CAATAATTCAATGGATAAATCTGTCTTTGAAATATAACATTATGCTGCCTGGATGATATGCATAT
TAAACATATTTGGAAAAGTGGAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

[illegible]

Signal peptide:
amino acids 1-16

FIGURE 97

GCTCAAGTGGCCCTGCCTTGCCCCACCCAGCCAGCCTGGCCAGAGCCCCCTGGAGAAGGAGCTCT
CTTCTTGCTTGGCAGCTGGACCAAGGGAGCCAGTCTTGGGCGCTGGAGGGCCTGTCTTGACCATG
GTCCCTGCCTGGCTGTGGCTGCTTTGTGTCTCCGTCCCCCAGGCTCTCCCCAAGGCCAGCCTGC
AGAGCTGTCTGTGGAAGTTCAGAAAACTATGGTGGAAATTTCCCTTTATACCTGACCAAGTTGC
CGCTGCCCCGTGAGGGGGCTGAAGGCCAGATCGTGCTGTGTCAGGGGACTCAGGCAAGGCAACTGAG
GGCCCATTTGCTATGGATCCAGATTCTGGCTTCCTGCTGGTGACCAGGGCCCTGGACCGAGAGGA
GCAGGCAGAGTACCAGCTACAGGTCACCCCTGGAGATGCAGGATGGACATGTCTTGTGGGTCCAC
AGCCTGTGCTTGTGCACGTGAAGGATGAGAATGACCAGGTGCCCCATTTCTCTCAAGCCATCTAC
AGAGCTCGGCTGAGCCGGGGTACCAGGCTGGCATCCCCCTTCCTCTTCCCTGAGGCTTCAGACCG
GGATGAGCCAGGCACAGCCAACCTCGGATCTTCGATTCCACATCCTGAGCCAGGCTCCAGCCCAGC
CTTCCCCAGACATGTTCCAGCTGGAGCCTCGGCTGGGGGCTCTGGCCCTCAGCCCCAAGGGGAGC
ACCAGCCTTGACCACGCCCTGGAGAGGACCTACCAGCTGTTGGTACAGGTCAAGGACATGGGTGA
CCAGGCCTCAGGCCACCAGGCCACTGCCACCGTGGAAAGTCTCCATCATAGAGAGCACCTGGGTGT
CCCTAGAGCCTATCCACCTGGCAGAGAATCTCAAAGTCCCTATACCCGACACATGGGCCAGGTA
CACTGGAGTGGGGGTGATGTGCACTATCACCTGGAGAGCCATCCCCGGGACCCCTTTGAAGTGAA
TGCAGAGGGAAACCTCTACGTGACCAGAGAGCTGGACAGAGAAGCCAGGCTGAGTACCTGCTCC
AGGTGCGGGCTCAGAATTTCCATGGCGAGGACTATGCGGGCCCTCTGGAGCTGCACGTGCTGGTG
ATGGATGAGAATGACAACGTGCCTATCTGCCCTCCCCGTGACCCACAGTCAGCATCCCTGAGCT
CAGTCCACCAGGTACTGAAGTGACTAGACTGTGTCAGCAGAGGATGCAGATGCCCCGGCTCCCCCA
ATTTCCACGTTGTGTATCAGCTCCTGAGCCCTGAGCCTGAGGATGGGGTAGAGGGGAGAGCCTTC
CAGGTGGACCCCACTTCAGGCAGTGTGACGCTGGGGGTGCTCCCACTCCGAGCAGGCCAGAACAT
CCTGCTTCTGGTGCTGGCCATGGACCTGGCAGGCGCAGAGGGTGGCTTCAGCAGCACGTGTGAAG
TCGAAGTCGCAGTCACAGATATCAATGATCACGCCCTGAGTTCATCACTTCCAGATTGGGCCT
ATAAGCCTCCCTGAGGATGTGGAGCCCGGACTCTGGTGGCCATGCTAACAGCCATTGATGCTGA
CCTCGAGCCCGCCTTCCGCCTCATGGATTTTGCCATTGAGAGGGGAGACACAGAAGGGACTTTTG
GCCTGGATTGGGAGCCAGACTCTGGGCATGTTAGACTCAGACTCTGCAAGAACCTCAGTTATGAG
GCAGCTCCAAGTCATGAGGTGGTGGTGGTGGTGCAGAGTGTGGCGAAGCTGGTGGGGCCAGGCC
AGGCCTGGAGCCACCGCCACGGTGACTGTGCTAGTGGAGAGAGTGTGACACCCCCCAAGTTGG
ACCAGGAGAGCTACGAGGCCAGTGTCCCCATCAGTGCCCCAGCCGGCTCTTTCCTGCTGACCATC
CAGCCCTCCGACCCCATCAGCCGAACCTCAGGTTCTCCCTAGTCAATGACTCAGAGGGCTGGCT
CTGCATTGAGAAATTTCTCCGGGGAGGTGCACACCGCCAGTCCCTGCAGGGCGCCAGCCTGGGG
ACACCTACACGGTGCTTGTGGAGGCCAGGATACAGCCCTGACTCTTGCCCTGTGCCCTCCCAA
TACCTCTGCACACCCCGCCAAGACCATGGCTTGATCGTGAGTGGACCCAGCAAGGACCCCGATCT
GGCCAGTGGGCACGGTCCCTACAGCTTACCCCTTGGTCCCAACCCACGGTGCAACGGGATTGGC
GCCTCCAGACTCTCAATGGTTCCCATGCCTACCTCACCTTGGCCCTGCATTGGGTGGAGCCAGT
GAACACATAATCCCGTGGTGGTGCAGCCACAATGCCAGATGTGGCAGCTCCTGGTTCCAGTGAT
CGTGTGTGCTGCAACGTGGAGGGGCAGTGCATGCGCAAGGTGGGCCGATGAAGGGCATGCCCA
CGAAGCTGTGCGCAGTGGGCATCCTTGAGGCACCTGGTAGCAATAGGAATCTTCTCATCCTC
ATTTTCACCCACTGGACCATGTCAAGGAAGAAGGACCCGGATCAACCAGCAGACAGCGTGCCCT
GAAGGCGACTGTCTGAATGGCCAGGCAGCTTAGCTGGGAGCTTGGCCTCTGGCTCCATCTGAG
TCCCCTGGGAGAGAGCCAGCACCCAAGATCCAGCAGGGGACAGGACAGAGTAGAAGCCCTCCA
TCTGCCCTGGGGTGGAGGCACCATCACCATCACCAGGCATGTCTGCAGAGCCTGGACACCAACTT
TATGGACTGCCCATGGGAGTGCTCCAAATGTGAGGGTGTTCGCCAATAATAAAGCCCCAGAGAA
CTGGGCTGGGCCCTATGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG

FIGURE 98

MVPAWLWLLCVSPQALPKAQAELSVEVPENYGGNFPLYLTKLPLPREGAEGQIVLSGDSGKAT
 EGPFAMDPDSGFLLVTRALDREEQAEYQLQVTLEMQDGHVLWGPQPVLVHVKDENDQVPHFSQAI
 YRARLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQSPDMFQLEPRLGALALSPKG
 STSLDHALERTYQLLVQVKMDGDQASGHQATATVEVSI IESTWVSLEPIHLAENLKVLYPHHMAQ
 VHWSGGDVHYHLESHPPGPFVNAEGNLYVTRELDREAQAEYLLQVRAQNSHGEDIAAPLELHVL
 VMDENDNVPICPPRDPTVSIPELSPPGTEVTRLSAEDADAPGSPNSHVYQLLSPEPEDGVEGRA
 FQVDPTSGSVTLGVLPLRAGQNILLVLAMDLAGAEGGFSSTCEVEVAVTDINDHAPEFITSQIG
 PISLPEDVEPGTLVAMLTADADLEPAFRIMDFAIERGDTEGTFGLDWEPPDSGHVRLRLCKNLSY
 EAAPSHEVVVVVQSVAKLVGPGPGGATATVTVLVERVMPPPKLDQESYEASVPISAPAGSFLLT
 IQPSDPISRTLRFSLVNDSEGWLCEIKFSGEVHTAQSLQGAQPGDTYTVLVEAQDTALT LAPVPS
 QYLCTPRQDHGLIVSGPSKDPDLASGHGPYSFTLGPNPTVQRDWRLQTLNGSHAYLTALHWVEP
 REHIIPVVVSHNAQMWQLLVRVIVCRCNVEGQCMRKVGRMKGMPTKLSAVGILVGTILVAIGIFLI
 LIFTHWTMSRKKDPDQPADSVPLKATV

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 762-784

FIGURE 99

GGCTGACCGTGCTACATTGCCTGGAGGAAGCCTAAGGAACCCAGGCATCCAGCTGCCCCAGCCTG
AGTCCAAGATTCTTCCCAGGAACACAAACGTAGGAGACCCACGCTCCTGGAAGCACCAGCCTTTA
TCTCTTACCTTCAAGTCCCCCTTTCTCAAGAATCCTCTGTTCTTTGCCCTCTAAAGTCTTGGTAC
ATCTAGGACCCAGGCATCTTGCTTTCCAGCCACAAAGAGACAGATGAAGATGCAGAAAGGAAATG
TTCTCCTTATGTTTGGTCTACTATTGCATTAGAAGCTGCAACAAATTCCAATGAGACTAGCACC
TCTGCCAACACTGGATCCAGTGTGATCTCCAGTGGAGCCAGCACAGCCACCAACTCTGGGTCCAG
TGTGACCTCCAGTGGGGTCCAGCACAGCCACCATCTCAGGGTCCAGCGTGACCTCCAATGGGGTCA
GCATAGTCACCAACTCTGAGTTCATACAACCTCCAGTGGGATCAGCACAGCCACCAACTCTGAG
TTCAGCACAGCGTCCAGTGGGATCAGCATAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGG
GGCCAGCACAGCCACCAACTCTGAGTCCAGCACACCCTCCAGTGGGGCCAGCACAGTCCCAACT
CTGGGTCCAGTGTGACCTCCAGTGGAGCCAGCACTGCCACCAACTCTGAGTCCAGCACAGTGTCC
AGTAGGGCCAGCACTGCCACCAACTCTGAGTCTAGCACACTCTCCAGTGGGGCCAGCACAGCCAC
CAACTCTGACTCCAGCACAACTCCAGTGGGGCTAGCACAGCCACCAACTCTGAGTCCAGCACAA
CCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGTGTCCAGTAGGGCCAGCACT
GCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAG
AACGACCTCCAATGGGGCTGGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGGGGCCA
GCACAGCCACCAACTCTGACTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAG
TCCAGCACGACCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGG
GGCTAGCACAGCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACT
CTGAGTCCAGCACAGTGTCCAGTGGGATCAGCACAGTACCAATTCTGAGTCCAGCACACCCTCC
AGTGGGGCCAACACAGCCACCAACTCTGAGTCCAGTACGACCTCCAGTGGGGCCAACACAGCCAC
CAACTCTGAGTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAGTCCAGCACAA
CCTCCAGTGGGGTCCAGCACAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCTAGCAC
GCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCTAG
CACAGTGTCCAGTGGGATCAGCACAGTACCAATTCTGAGTCCAGCACAACTCCAGTGGGGCCA
ACACAGCCACCAACTCTGGGTCCAGTGTGACCTCTGCAGGCTCTGGAACAGCAGCTCTGACTGGA
ATGCACACAACTTCCCATAGTGCATCTACTGCAGTGAAGTGAAGCAAGCCTGGTGGGTCCCTGGT
GCCGTGGGAAATCTTCTCATCACCTGGTCTCGGTTGTGGCGGCCGTGGGGCTCTTTGCTGGGC
TCTTCTTCTGTGTGAGAAACAGCCTGTCCCTGAGAAACACCTTTAACACAGCTGTCTACCACTCT
CATGGCCTCAACCATGGCCTTGGTCCAGGCCCTGGAGGGAATCATGGAGCCCCCACAGGCCAG
GTGGAGTCTTAAGTGGTCTGGAGGAGACCAATCATCGATAGCCATGGAGATGAGCGGGAGGA
ACAGCGGGGCCCTGAGCAGCCCCGGAAGCAAGTGCCGCATTCTTCAGGAAGGAAGAGACCTGGGCA
CCCAAGACCTGGTTCTCTTTTATTCATCCAGGAGACCCCTCCAGCTTTGTTTGAGATCCTGAA
AATCTTGAAGAAGGTATTCTTACCTTTCTTGCCTTTACAGACACTGGAAAGAGAATACTATAT
TGCTCATTTAGCTAAGAAATAAATACATCTCATCTAACACACACGACAAAGAGAAGCTGTGCTTG
CCCCGGGGTGGGTATCTAGCTCTGAGATGAAGTCAAGTTATAGGAGAAAACCTCCATGCTGGACTC
CATCTGGCATTCAAAATCTCCACAGTAAATCCAAAGACCTCAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 100

MMQKGNVLLMFGLLLHLEATNSNETSTSANTGSSVSSGASTATNSGSSVTSSGVSTATISGS
SVTSNGVSIIVTNSEFHTTSSGISTATNSEFSTASSGISIATNESSTTSSGASTATNESSTPSS
GASTVTNSGSSVTSSGASTATNESSTVSSRASTATNESSTLSSGASTATNSDSSTTSSGASTA
TNSESSTTSSGASTATNESSTVSSRASTATNESSTTSSGASTATNESRTTNGAGTATNES
STTSSGASTATNSDSSTVSSGASTATNESSTTSSGASTATNESSTTSSGASTATNSDSSTSS
GAGTATNESSTVSSGISTVTNESSTPSSGANTATNESSTTSSGANTATNESSTVSSGASTA
TNSESSTTSSGVSTATNESSTTSSGASTATNSDSSTTSEASTATNESSTVSSGISTVTNES
STTSSGANTATNSGSSVTSAGSGTAALTGMHTTSHSASTAVSEAKPGGSLVPWEIFLITLVSVA
AVGLFAGLFFCVRNSLSLRNTFNTAVYHPHGLNHGLGPGPGGNHGAPHRPRWSPNWFWRPVSII
AMEMSGRNSGP

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 510-532

FIGURE 101

GGCCGGACGCCTCCGCGTTACGGGATGAATTAACGGCGGGTTCCGCACGGAGGTTGTGACCCCTA
CGGAGCCCCAGCTTGCCACGCACCCCACTCGGCGTCGCGCGGCGTGCCCTGCTTGTACAGGTG
GGAGGCTGGAACATCAGGCTGAAAAACAGAGTGGGTACTCTCTTCTGGGAAGCTGGCAACAAAT
GGATGATGTGATATATGCATTCCAGGGGAAGGGAAATTGTGGTGCTTCTGAACCCATGGTCAATT
AACGAGGCAGTTTCTAGCTACTGCACGTACTTCATAAAGCAGGACTCTAAAAGCTTTGGAATCAT
GGTGTGATGGAAGGGATTTACTTTTATACTGACTCTGTTTTGGGGAAGCTTTTTTGGGAAGCATTT
TCATGCTGAGTCCCTTTTTACCTTTGATGTTTGTAAACCCATCTTGGTATCGCTGGATCAACAAC
CGCCTTGTGGCAACATGGCTCACCTACCTGTGGCATTATTGGAGACCATGTTTGGTGTAAAAGT
GATTATAACTGGGGATGCATTTGTTCTGAGAAAGAAGTGTCAATTATCATGAACCATCGGACAA
GAATGGACTGGATGTTCTGTGGAATTGCCTGATGCGATATAGCTACCTCAGATTGGAGAAAATT
TGCCTCAAAGCGAGTCTCAAAGGTGTTCTGGATTGTTTGGGCCATGCAGGCTGCTGCCTATAT
CTTCATTCATAGGAAATGGAAGGATGACAAGAGCCATTTGAAGACATGATTGATTACTTTTGTG
ATATTCACGAACCACTTCAACTCCTCATATCCCAGAAGGGACTGATCTCAGAGAAAACAGCAAG
TCTCGAAGTAATGCATTTGCTGAAAAAATGGACTTCAGAAATATGAATATGTTTTACATCCAAG
AACTACAGGCTTTACTTTTGTGGTAGACCGTCTAAGAGAAGGTAAGAACCTTGATGCTGTCCATG
ATATCACTGTGGCGTATCCTCACAACATTCCTCAATCAGAGAAGCACCTCCTCCAAGGAGACTTT
CCCAGGGAAATCCACTTTCACGTCCACCGGTATCCAATAGACACCTCCCCACATCCAAGGAGGA
CCTTCAACTCTGGTGCCACAAACGGTGGGAAGAGAAAGAAGAGAGGCTGCGTTCCTTCTATCAAG
GGGAGAAGAATTTTTATTTTACCGGACAGAGTGTCAATCCACCTTGCAAGTCTGAACTCAGGGTC
CTTGTGGTCAAATTGCTCTCTATACTGTATTGGACCCTGTTTCAAGCCCTGCAATGTGCCTACTCAT
ATATTTGTACAGTCTTGTAAAGTGGTATTTTATAATCACCATTGTAATCTTTGTGCTGCAAGAGA
GAATATTTGGTGGACTGGAGATCATAGAACTTGCATGTTACCGACTTTTACACAAACAGCCACAT
TTAAATTCAAAGAAAAATGAGTAAGATTATAAGGTTTGCCATGTGAAAACCTAGAGCATATTTTG
GAAATGTTCTAAACCTTTCTAAGCTCAGATGCATTTTTGCATGACTATGTGCAATATTTCTTACT
GCCATCATTATTTGTAAAGATATTTGCACCTAATTTTGTGGGAAAAATATTGCTACAAATTTT
TTAATCTCTGAATGTAATTCGATACTGTGTACATAGCAGGGAGTGATCGGGGTGAAATAACTT
GGGCCAGAATATTATTAAACAATCATCAGGCTTTTAA

FIGURE 102

MHSRGREIVVLLNPWSINEAVSSYCTYFIKQDSKSFIMVSWKGIYFILTLFWGSFFGSIIFMLSP
FLPLMFVNPSWYRWINNRLVATWLTLPVALLETMFGVKVIITGDAFVPGERSVIIMNHRTRMDWM
FLWNCLMRYSYLRLEKICLKASLKGVPFGFWAMQAAAYIFIHRKWKDDKSHFEDMIDYFCDIHEP
LQLLI FPEGTDLTENSKSRSNFAEKNGLQKYEYVLHPRTTGFTFVVDRLREGKNLDAVHDITVA
YPHNIPQSEKHLQGDFFREIHFHVHRYPIDTLPTSKEDLQLWCHKRWEEKEERLSFYQGEKNF
YFTGQSVIPPCKSELRLVLVVKLLSILYWTLFSPAMCLLIYLYSLVKWYFIITIVIFVLQERIFGG
LEIIELACYRLLHKQPHLNSKKNE

Important features of the protein:

Signal peptide:

amino acids 1-22

Transmembrane domains:

amino acids 44-63, 90-108, 354-377

FIGURE 103

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCTCCACGGCTCCTGCGCTGAGACAGCTGGCCTGACC
TCCAAATCATCCATCCACCCCTGCTGTCATCTGTTTTCATAGTGTGAGATCAACCCACAGGAATA
TCCATGGCTTTTTGTGCTCATTTTTGGTTCTCAGTTTCTACGAGCTGGTGTGACAGTGGCAAGT
CACTGGACCGGGCAAGTTTGTCCAGGCCTTGGTGGGGGAGGACGCCGTGTTCTCTGCTCCCTCT
TTCCTGAGACCAAGTGCAGAGGCTATGGAAGTGCCTTCTTCAGGAATCAGTTCCATGCTGTGGTC
CACCTCTACAGAGATGGGGAAGACTGGGAATCTAAGCAGATGCCACAGTATCGAGGGAGAATGA
GTTTGTGAAGGACTCCATTGACAGGGGGCGTGTCTCTAAGGCTAAAAACATCACTCCCTCGG
ACATCGGCCTGTATGGGTGCTGGTTCAGTTCCAGATTTACGATGAGGAGGCCACCTGGGAGCTG
CGGGTGGCAGCACTGGGCTCACTTCCTCTCATTTCATCGTGGGATATGTTGACGGAGGTATCCA
GTTACTCTGCCTGTCTCAGGCTGGTTCCCCAGCCACAGCCAAGTGGAAAGGTCCACAAGGAC
AGGATTTGTCTTCAGACTCCAGAGCAAATGCAGATGGGTACAGCCTGTATGATGTGGAGATCTCC
ATTATAGTCCAGGAAATGCTGGGAGCATATTGTGTTCCATCCACCTTGCTGAGCAGAGTATGA
GGTGAATCCAAGGTATTGATAGGAGAGACGTTTTTCCAGCCCTCACCTTGGCGCCTGGCTTCTA
TTTTACTCGGGTTACTCTGTGGTGCCCTGTGTGGTGTGTGTCATGGGGATGATAATTGTTTTCTTC
AAATCCAAAGGGAAATCCAGGCGGAAGTGGACTGGAGAAGAAAGCACGGACAGGCAGAAATTGAG
AGACGCGCGGAAACACGCAGTGGAGGTGACTCTGGATCCAGAGACGGCTCACCGAAGCTCTGCG
TTTCTGATCTGAAACTGTAACCCATAGAAAAGCTCCCCAGGAGGTGCCTCACTCTGAGAAGAGA
TTTACAAGGAAGAGTGTGGTGGCTTCTCAGGGTTTCCAAGCAGGAGACATTACTGGGAGGTGGA
CGTGGGACAAAATGTAGGGTGGTATGTGGGAGTGTGTCGGGATGACGTAGACAGGGGGAAGAAC
ATGTGACTTTGTCTCCCAACAATGGGTATTGGGTCTCAGACTGACAACAGAACATTTGTATTTT
ACATTCAATCCCCATTTTATCAGCCTCCCCCCCAGCACCCCTCCTACACGAGTAGGGGTCTTCCT
GGACTATGAGGGTGGGACCATCTCCTTCTTCAATACAAATGACCAGTCCCTTATTTATACCTGC
TGACATGTGAGTTTGAAGGCTTGTGAGACCCCTATATCCAGCATGCGATGTATGACGAGGAAAAG
GGGACTCCCATATTCATATGTCCAGTGTCTGGGGATGAGACAGAGAAGACCCCTGCTTAAAGGGC
CCCACACCACAGACCAGACACAGCCAAGGGAGAGTGTCTCCGACAGGTGGCCCCAGCTTCTCT
CCGGAGCCTGCGCACAGAGAGTCAAGCCCCCACTCTCCTTTAGGGAGCTGAGGTTCTTCTGCCC
TGAGCCCTGCAGCAGCGGCAGTCAAGCTTCCAGATGAGGGGGGATTGGCCTGACCCGTGGGGAG
TCAGAAGCCATGGCTGCCCTGAAGTGGGGACGGAATAGACTCACATTAGGTTTAGTTTGTGAAAA
CTCCATCCAGCTAAGCGATCTTGAACAAGTCACAACCTCCCAGGCTCCTCATTGCTAGTCACGG
ACAGTGATTCTGCTCACAGGTGAAGATTAAAGAGACAACGAATGTGAATCATGCTTGCAGGTT
TGAGGGCACAGTGTGCTAATGATGTGTTTTTATATTATACATTTTCCACCATAAACTCTGTT
TGCTTATTCACATTAATTTACTTTTCTCTATACCAAATCACCCATGGAATAGTTATTGAACACC
TGCTTTGTGAGGCTCAAAGAATAAAGAGGAGGTAGGATTTTTCTACTGATTCTATAAGCCCAGCAT
TACCTGATACCAAACAGGCAAAGAAAACAGAGAAGAGGAAGGAAAACACAGGTCCATATCC
CTCATTAACACAGACACAAAAATTTCTAAATAAAATTTTAAACAAATTAACATAAACAATATATTTA
AAGATGATATATAACTACTCAGTGTGGTTTGTCCACAAATGCAGAGTTGGTTTAATATTTAAAT
ATCAACCAGTGTAATTCAGCACATTAATAAAGTAAAAAGAAAACCATAAAAA

FIGURE 104

MAFVLILVLSFYELVSGQWQVTGPGKFVQALVGEDAVFSCSLFPETSAEAMEVRFFRNQFHAVVH
LYRDGEDWESKQMPQYRGRTEFVKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEEATWELR
VAALGSLPLISIVGYVDGGIQLLCLSSGWFPQPTAKWKGPQGDLSSDSRANADGYSLYDVEISI
IVQENAGSILCSIHLAEQSHEVESKVLIGETFFQPSPWRLASILLGLLCGALCGVVMGMIIVFFK
SKGKIQAELDWRKKGQAELRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAPQEVPHSEKRF
TRKSVVASQGFQAGRHYWEVDVGQNVGWYVGCRDDVDRGKNNVTLSPNNGYWVLRLTTEHLYFT
FNPHFISLPPSTPPTRVGVFLDYEGGTISFNTNDQSLIYTLTLCQFEGLLRPYIQHAMYDEEKG
TPIFICPVSWG

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 131-150, 235-259

FIGURE 105

CCTTCACAGGACTCTTCATTGCTGGTTGGCAATGATGTATCGGCCAGATGTGGTGAGGGCTAGGAAAAGAG
 TTTGTTGGGAACCTGGGTTATCGGCCTCGTCATCTTCATATCCCTGATTGTCTGGCAGTGTGCATTGGA
 CTCACTGTTTCATTATGTGAGATATAATCAAAAGAAGACCTACAATTACTATAGCACATTGTTCATTACAAAC
 TGACAAACTATATGCTGAGTTTGGCAGAGAGGCTTCTAACAATTTTACAGAAATGAGCCAGAGACTTGAAT
 CAATGGTGAAAATGCATTTTATAAATCTCCATTAAGGGAAGAATTTGTCAAGTCTCAGGTTATCAAGTTC
 AGTCAACAGAAGCATGGAGTGTGGCTCATATGCTGTTGATTGTAGATTTCACTCTACTGAGGATCCTGA
 AACTGTAGATAAAATTGTTCAACTTGTTTTACATGAAAAGCTGCAAGATGCTGTAGGACCCCCATAAGTAG
 ATCCTCACTCAGTTAAAATTAAAAAATCAACAAGACAGAAACAGACAGCTATCTAAACCATTGCTGCCGA
 ACACGAAGAAGTAAACTCTAGGTGAGAGTCTCAGGATCGTTGGTGGGACAGAAGTAGAAGAGGGTGAATG
 GCCCTGGCAGGCTAGCCTGCAGTGGGATGGGAGTCACTCGCTGTGGAGCAACCTTAATTAATGCCACATGGC
 TTGTGAGTGTCTCACTGTTTTTACAACATATAAGAACCCTGCCAGATGGACTGCTTCTTTGGAGTAACA
 ATAAACCTTCGAAAATGAAACGGGTCTCCGAGAATAATTGTCCATGAAAATACAAACACCCATCACA
 TGACTATGATATTTCTCTTGACAGAGCTTTCTAGCCCTGTTCCCTACACAAATGCAGTACATAGAGTTTGT
 TCCCTGATGCATCCTATGAGTTTCAACCAGGTGATGTGATGTTGTGACAGGATTTGGAGCACTGAAAAAT
 GATGGTTACAGTCAAAATCATCTTCGACAAGCACAGGTGACTCTCATAGACGCTACAACCTGCAATGAACC
 TCAAGCTTACAATGACGCCATAACTCCTAGAATGTTATGTGCTGGCTCCTTAGAAGGAAAAACAGATGCAT
 GCCAGGGTGACTCTGGAGGACCACTGGTTAGTTCAGATGCTAGAGATATCTGTTACCTTGCTGGAATAGTG
 AGCTGGGGAGATGAATGTGCGAAACCAACAAGCCTGGTGTATATACTAGAGTTACGGCCTTGCGGGACTG
 GATTACTTCAAAAACCTGGTATCTAAAGAGACAAAAGCCTCATGGAACAGATAACATTTTTTTTTGTTTTTG
 GGTGTGGAGGCCATTTTTAGAGATACAGAATTGGAGAAGACTTGCAAAACAGCTAGATTTGACTGATCTCA
 ATAAACTGTTTGCTTGATGCATGTATTTCTTCCAGCTCTGTTCCGCACGTAAGCATCCTGCTTCTGCCA
 GATCAACTCTGTCTGTGAGCAATAGTTGAACTTTATGTACATAGAGAAATAGATAATACAATATTAC
 ATTACAGCCTGTATTCAATTGTTCTCTAGAAGTTTTGTGAGAATTTTGACTTGTGACATAAATTTGTAAT
 GCATATATACAATTTGAAGCACTCCTTTTCTTCAGTTCCTCAGCTCCTCTCATTTCAGCAAATATCCATTT
 TCAAGGTGCAGAACAGGACTGAAGAAAAATATAAGAAGAAAAAATCCCCACATTTTATTGGCACAGAA
 AAGTATTAGGTGTTTTCTTAGTGGAATATTAGAAATGATCATATTATTATGAAAGGTCAAGCAAAGACA
 GCAGAATACCAATCACTTCATCATTTAGGAAGTATGGGAACCTAAGTTAAGGAAGTCCAGAAAGAAGCCAAG
 ATATATCCTTATTTTCATTTCCAAACAACACTACTATGATAAATGTGAAGAAGATTCTGTTTTTTGTGACCT
 ATAATAATTATACAACTTCATGCAATGTACTTGTCTAAGCAAATTAAAGCAAATATTTATTTAACATTG
 TTACTGAGGATGTCAACATATAACAATAAAATATAAATCACCCA

FIGURE 106

MMYRPDVVRARKRVCWEPWVIGLVIFISLIVLAVCIGLTVHYVRYNQKKTNYNSTLSFTTDKLY
AEFGREASNNFTEMSQRLESMVKNAFYKSPLEEFVKSQVIKFSQQKHGVLAHMLLICRFHSTED
PETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTETDSYLNHCCGTRRSKTLGQSLRIVGG
TEVEEGEWPWQASLQWDGSHRCGATLINATWLVSAHCFTTYKNPARWTASFGVTIKPSKMKRGL
RRIIVHEKYKHPSHDYDISLAEISSPVYPTNAVHRVCLPDASYEFQPGDVMFVTGFGALKNDGYS
QNHLRQAQVTLIDATTCNEPQAYNDAITPRMLCAGSLEGKTDACQGDSGGPLVSSDARDIWYLAG
IVSWGDECAKPNKPGVYTRVTALRDWITSKTGI

Transmembrane domain:

amino acids 21-40 (type II)

FIGURE 107

AGAGAAAGAAGCGTCTCCAGCTGAAGCCAATGCAGCCCTCCGGCTCTCCGCGAAGAAGTTCCTG
CCCCGATGAGCCCCCGCGTGCCTCCCCGACTATCCCCAGGCGGGCGTGGGGCACCAGGGCCCAGC
GCCGACGATCGCTGCCGTTTTGCCCCTGGGAGTAGGATGTGGTGAAAGGATGGGGCTTCTCCCTT
ACGGGGCTCACAAATGGCCAGAGAAGATTCGCTGAAGTGTCTGCGCTGCCTGCTCTACGCCCTCAA
TCTGCTCTTTTGGTTAATGTCCATCAGTGTGTTGGCAGTTTCTGCTTGGATGAGGGACTACCTAA
ATAATGTTCTCACTTTAACTGCAGAAACGAGGGTAGAGGAAGCAGTCATTTGACTTACTTTCCT
GTGGTTCATCCGGTCATGATTGCTGTTTGCTGTTTCCTTATCATTGTGGGGATGTTAGGATATTG
TGGAACGGTGAAAGAAATCTGTTGCTTCTTGCATGGTACTTTGGAAGTTTGCTTGTCAATTTCT
GTGTAGAACTGGCTTGTGGCGTTTGGACATATGAACAGGAACCTATGGTTCCAGTACAATGGTCA
GATATGGTCACTTTGAAAGCCAGGATGACAAATTATGGATTACCTAGATATCGGTGGCTTACTCA
TGCTTGAATTTTTTTCAGAGAGAGTTTAAAGTCTGTGGAGTAGTATATTTCACTGACTGGTTGG
AAATGACAGAGATGGACTGGCCCCAGATTCTGCTGTGTTAGAGAATTCAGGATGTTCCAAA
CAGGCCCACCAGGAAGATCTCAGTGACCTTTATCAAGAGGGTGTGGGAAGAAAATGTATTCCTT
TTTGAGAGGAACCAACAACACTGCAGGTGCTGAGGTCTTCTGGGAATCTCCATTGGGGTGACACAAA
TCCTGGCCATGATTCTCACCATTACTCTGCTCTGGGCTCTGTATTATGATAGAAGGGAGCCTGGG
ACAGACCAATGATGTCCTTGAAGAATGACAACCTCTCAGCACCTGTCTATGTCCTTCAGTAGAACT
GTTGAACCAAGCCTGTCAAGAATCTTTGAACACACATCCATGGCAAACAGCTTTAATACACACT
TTGAGATGGAGGAGTTATATAAAGAAATGTACAGAGAAGAAAACCAAACTTGTTTTATTGGACT
TGTGAATTTTTGAGTACATACATATGTGTTTCAGAAATATGTAGAAATAAAAATGTTGCCATAAAA
TAACACCTAAGCATATACTATTCTATGCTTTAAATGAGGATGGAAAAGTTTCATGTCTAAGTC
ACCACCTGGACAATAATTGATGCCCTTAAATGCTGAAGACAGATGTCTATCCCACTGTGTAGCC
TGTGTATGACTTTTACTGAACACAGTTATGTTTTGAGGCAGCATGGTTTGATTAGCATTTCCGCA
TCCATGCAAACGAGTCACATATGGTGGGACTGGAGCCATAGTAAAGGTTGATTACTTCTACCAA
CTAGTATATAAAGTACTAATTAATGCTAACATAGGAAGTTAGAAAATACTAATAACTTTTATTA
CTCAGCGATCTATTCTCTGATGCTAAATAAATTATATATCAGAAAACCTTCAATATTGGTGACT
ACCTAAATGTGATTTTTGCTGGTTACTAAAATATCTTACCCTTAAAAGAGCAAGCTAACACAT
TGTCTTAAGCTGATCAGGGATTTTTGTATATAAGTCTGTGTTAAATCTGTATAAATTCAGTCGAT
TTCAGTTCTGATAATGTTAAGAATAACCATTATGAAAAGGAAAATTTGTCCTGTATAGCATCATT
ATTTTTAGCCTTTCCGTGTTAATAAAGCTTTACTATTCTGTCTGGGCTTATATTACACATATAAC
TGTTATTTAAATACTTAACCACTAATTTTGAATAATACCAGTGTGATACATAGGAATCATTATTC
AGAATGTAGTCTGGTCTTTAGGAAGTATTAATAAGAAAATTTGCACATAACTTAGTTGATTTCAGA
AAGGACTTGTATGCTGTTTTTCTCCCAAATGAAGACTCTTTTGGACACTAAACACTTTTTAAAAA
GCTTATCTTTGCCTTCTCCAAACAAGAAGCAATAGTCTCCAAGTCAATATAAATCTACAGAAAA
TAGTGTCTTTTTCTCCAGAAAAATGCTTGTGAGAATCATTAACACATGTGACAATTTAGAGATT
CTTTGTTTTATTTCACTGATTAATATACTGTGGCAAATTACACAGATTATTAAATTTTTTACAA
GAGTATAGTATATTTATTTGAAATGGGAAAAGTGCATTTTACTGTATTTTGTGTATTTTGTGTTAT
TTCTCAGAATATGGAAGAAAATTAATATGTGTCAATAAATATTTCTAGAGAGTAA

FIGURE 108

MAREDSVKCLRCLLYALNLLFWLMSISVLAVSAWMRDYLNNVLTTLTAETRVEEAVILTYFPVVHP
 VMIAVCCFLIIVGMLGYCGTVKRNLLLLAWYFGSLLVIFCVELACGVWTYEQELMVPVQWSDMVT
 LKARMTNYGLPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTMDWPPDSCCVREFPGCSKQAHQ
 EDLSDLYQEGCGKKMYSFLRGTKQLQVLRFLGISIGVTQILAMILTITLLWALYYDRREPGTDQM
 MSLKNDNSQHLSCPSVELLKPSLSRIFEHTSMANSFNTHFEMEEL

Signal peptide:

amino acids 1-33

Transmembrane domains:

amino acids 12-35, 57-86, 94-114, 226-248

FIGURE 109

CCAAGGCCAGAGCTGTGGACACCTTATCCCACTCATCCTCTTCTCTGATAAAGCCCCCTACCAGTGCT
GATAAAGTCTTTCTCGTGAGAGCCTAGAGGCCCTTAAAAAAGTGCTTGAAAGAGAAGGGGACAAAGGAACA
CCAGTATTAAGAGGATTTCCAGTGTTTCTGGCAGTTGGTCCAGAAGGATGCCCTCCATTCTGCTTCTCACCTG
CCTCTTCATCACAGGCACCTCCGTGTCAACCGTGGCCCTAGATCCTTGTTCTGCTTACATCAGCCTGAATGAGC
CCTGGAGGAACACTGACCACCAGTTGGATGAGTCTCAAGGTCTCTCTATGTGACAACCATGTGAATGGGAG
TGGTACCCTTACGGGCATGGCGGGAGATGCCATGCCTACCTTCTGCATACCAGAAAACCACTGTGGAACCCA
CGCACCTGTCTGGCTCAATGGCAGCCACCCCTAGAAGGCGACGGCATTGTGCAACCCAGGCTTGTGCCAGCT
TCAATGGGAAGTGTCTGTCTGGAACACCACGGTGGAAAGTCAAGGCTTGCCTGGAGGCTACTATGTGTATCGT
CTGACCAAGCCAGCGTCTGCTTCCACGTCTACTGTGGTCATTTTTATGACATCTGCGACGAGGACTGCCATGG
CAGCTGCTCAGATACCAGCGAGTGCACATGCCGTCCAGGAAGTGTGCTAGGCCCTGACAGGCAGACATGCTTTG
ATGAAAATGAATGTGAGCAAAACACGGTGGCTGCAGTGAGATCTGTCTGAACCTCAAAAACCTCTACCCTGT
GAGTGTGGGGTTGGCCGTGTGCTAAGAAGTGATGGCAAGACTTGTGAAGACGTTGAAGGATGCCACAATAACAA
TGGTGGCTGCAGCCACTCTTGCCTTGGATCTGAGAAAGGCTACCAGTGTGAATGTCCCCGGGGCCTGGTGTGT
CTGAGGATAACCACACTTGCCAAGTCCCTGTGTGTGCAAAATCAAATGCCATTGAAGTGAACATCCCCAGGGAG
CTGGTTGGTGGCCTGGAGCTCTTCTGACCAACACCTCTGCGGAGGAGTGTCCAACGGCACCCATGTCAACAT
CCTCTTCTCTCAAGACATGTGTGACAGTGGTGCATGTGGTGAATGACAAGATTGTGGCCAGCAACCTCGTGA
CAGGCTTACCAAGCAGACCCCGGGGAGCAGCGGGGACTTCATCATCCGAACCAAGAGCTGTGATCCCGGTG
ACCTGCCAGTTCACACGCTCTACACCATTTCTGAAGGATACGTTCCCAACCTTCGAAACTCCCCACTGGAAAT
CATGAGCCGAAATCATGGGATCTTCCCATTCACCTCTGGAGATCTTCAAGGACAATGAGTTTGAAGAGCCTTACC
GGGAAGCTCTGCCCCACCTCAAGCTTCTGTGACTCCCTCTACTTTGGCATTGAGCCCGTGGTGCACGTGAGCGGC
TTGGAAGCTTGGTGGAGAGCTGCTTTGCCACCCCACTCCCAAGATCCAGAGGCTCTGAAATACTACCTCAT
CCGGGATGGCTGTGTTTCAGATGACTCGGTAAAGCAGTACACATCCCGGGATCACCTAGCAAAGCACTTCCAGG
TCCCTGTCTTCAAGTTTGTGGGCAAGACCACAAGGAAGTGTCTTCTGCACTGCCGGGTTCTTGTCTGTGGAGTG
TTGGACGAGCGTTCCCGCTGTGCCCAGGGTTGCCACCGGCCAATGCGTCTGGGGCAGGAGGAGAGGACTCAGC
CGGTCTACAGGGCCAGACGCTAACAGGCGGCCGATCCGCATCGACTGGGAGGACTAGTTTCGTAGCCATACCTC
GAGTCCCTGCATTGGACGGCTCTGCTCTTTGGAGCTTCTCCCCCACCAGCTCTAAGAACATCTGCCAACAGC
TGGCTTCAGACTTCACACTGTGAGTTCAGACTCCAGCACCAACTCACTCTGATTCTGGTCCATTCACTGGGCA
CAGGTCACAGCACTGCTGAACAATGTGGCCTGGGTGGGGTTTCATCTTTCTAGGGTTGAAAATAAAGTGTCCA
CCCAGAAAGACACTACCCCATTTCCCTCATTTCTTTCTACACTAAATACCTCGTGTATGGTGAATCAGAC
CACAAAATCAGAAGCTGGGTATAATATTTCAAGTTACAAACCCTAGAAAAATTAAACAGTTACTGAAATTATGA
CTTAAATACCAATGACTCCTTAAATATGTAAATTATAGTTATACCTTGAATTTCAATTCAAATGCAGACTAA
TTATAGGGAATTTGGAAGTGTATCAATAAACAGTATATAATTTT

FIGURE 110

MPPFLLLTCLFITGTSVSPVALDPCSAYISLNPEWRNTDHQLDESQGPPLCDNHVNGEWYHFTGMAGDAMP
 TFCIPENHCGTHAPVWLNLSHPLEGDGIVQRQACASFNGNCCLWNTTVEVKACPGGYVYRLTKPSVCFHV
 YCGHFYDIDCEDCHGSCSDTSECTCAPGTVLGPDRQTCFDENECEQNNGGCSEICVNLKNSYRCECGVGRV
 LRS DGKTCEDVEGCHNNNGGCSHSLGSEKGYQCECPRLVLS DNHTCQVPVLCKSNAIEVNIPREL VGG
 LELFLTNTSCRGVSNNGTHVNILFSLKTCGTVVDVNDKIVASNLVTGLPKQTPGSSGDFIIRTSKLLIPVT
 CEFPRLYTISEGYVFNLRNSPLEIMSRNHGIFPFTLEIFKDNEFEOPYREALPTLKLRLDSLYFGIEPVVHV
 SGLESLVESCFATPTSKIDEVLKYYLIRDGCVSDDSVKQYTSRDHLAKHFQVPVFKFVKGDKHKEVFLHCRV
 LVCGVLDERSRCAQGCCHRRMRGAGGEDSAGLQGQTLTGGPPIRIDWED

Important features of the protein:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 89-93, 116-120, 259-263, 291-295, 299-303

Tyrosine kinase phosphorylation sites.

amino acids 411-418, 443-451

N-myristoylation sites.

amino acids 226-232, 233-239, 240-246, 252-258, 296-302, 300-306,
 522-528, 531-537

Aspartic acid and asparagine hydroxylation site.

amino acids 197-209

ZP domain proteins.

amino acids 431-457

Calcium-binding EGF-like proteins.

amino acids 191-212, 232-253

FIGURE 111

GAGAGAGGCAGCAGCTTGTCTCAGCGGACAAGGATGCTGGGCGTGAGGGACCAAGGCCTGCCCTGCACTCGG
 GCCTCCTCCAGCCAGTGCTGACCAGGGACTTCTGACCTGCTGGCCAGCCAGGACCTGTGTGGGAGGCCCT
 CCTGTGCTTGGGGTGACAATCTCAGCTCCAGGCTACAGGGAGACCGGGAGGATCACAGAGCCAGCATGT
 TACAGGATCCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAACCCCTGCGCAAACCCCGTATCCCC
 ATGGAGACCTTCAGAAAGGTGGGGATCCCCATCATATAGCACTACTGAGCCTGGCGAGTATCATCATGT
 GGTGTCCCTCATCAAGGTGATTCTGGATAAATACTACTTCCCTCTGCGGGCAGCCTCTCCACTTCATCCCGA
 GGAAGCAGCTGTGTGACGGAGAGCTGGACTGTCCCTTGGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTC
 CCCGAAGGGCCTGCAGTGGCAGTCCGCCCTCTCCAAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCAC
 AGGGAAGTGGTTCTCTGCCTGTTCGACAACCTTCACAGAAGCTCTCGCTGAGACAGCCTGTAGGCAGATGG
 GCTACAGCAGAGCTGTGGAGATTGGCCAGACCAGGATCTGGATGTTGTTGAAATCACAGAAACAGCCAG
 GAGCTTCGCATGCGGAACTCAAGTGGGCCCTGTCTCTCAGGCTCCCTGGTCTCCCTGCACTGTCTTGCTG
 TGGGAAGAGCCTGAAGACCCCCGTGTGGTGGGTGGGGAGGAGGCCCTCTGTGGATTCTTGGCCTTGGCAGG
 TCAGCATCCAGTACGACAAACAGCACGTCTGTGGAGGGAGCATCCTGGACCCCCACTGGGTCTCACGGCA
 GCCCACTGCTTCAGGAAACATACCGATGTGTTCAACTGGAAGGTGCGGGCAGGCTCAGACAAACTGGGCAG
 CTTCCCATCCCTGGCTGTGGCCAAGATCATCATATTGAATTCAACCCCATGTACCCCAAAGACAATGACA
 TCGCCCTCATGAAGCTGCAGTTCCCACTCACTTTCTCAGGCACAGTCAGGCCCATCTGTCTGCCCTTCTTT
 GATGAGGAGCTCACTCCAGCCACCCCACTCTGGATCATTGGATGGGGCTTTACGAAGCAGAATGGAGGAA
 GATGTCTGACATACTGTGTCAGGCGTCAGTCCAGGTCAATTGACAGCACACGGTGCAATGCAGACGATGCGT
 ACCAGGGGGAAGTCACCGAGAAGATGATGTGTGCAGGCATCCCGAAGGGGGTGTGGACACCTGCCAGGGT
 GACAGTGGTGGGCCCTGATGTACCAATCTGACCAGTGGCATGTGGTGGGCATCGTTAGCTGGGGCTATGG
 CTGCGGGGGCCCGAGCACCCCAAGGATATACACCAAGGTCTCAGCCTATCTCAACTGGATCTACAATGTCT
 GGAAGGCTGAGCTGTAATGTCTGCTGCCCTTTGCACTGCTGGGAGCCGCTTCCTTCTGCCCTGCCACCT
 GGGGATCCCCCAAAGTCAGACACAGAGCAAGAGTCCCTTGGGTACACCCCTCTGCCACAGCCTCAGCAT
 TTCTTGAGAGCAGCAAAGGGCCTCAATTCCCTGTAAGAGACCCCTCGCAGCCAGAGGCGCCAGAGGAAGTCA
 GCAGCCCTAGCTCGGCCACACTTGGTGCTCCCAGCATCCCAGGGAGAGACACAGCCCACTGAACAAGGTCT
 CAGGGGTATTGCTAAGCCAAGAAGGAACCTTTCCACACTACTGAATGGAAGCAGGCTGTCTTGTAAAGCC
 CAGATCACTGTGGGCTGGAGAGGAGAAGGAAAGGGTCTGCGCCAGCCCTGTCCGTCTTCACCCATCCCCAA
 GCCTACTAGAGCAAGAAACAGTTGTAATATAAAATGCACTGCCCTACTGTTGGTATGACTACCGTTACCT
 ACTGTTGTCAATTGTTATTACAGCTATGGCCACTATTATTAAAGAGCTGTGTAACATCTCTGGCAAAAAAA
 AAAA

FIGURE 112

MLQDPDSDQPLNSLDVKPLRKPRIPMETFRKVGIPIIIIALLSLASIIIVVVLIKVILDKYYFLCG
QPLHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLDSATGNWFSACFDN
FTEALAETACRQMGYSRAVEIGPDQDLDVVEITENSQELMRNSSGPCLSGSLVSLHCLACGKSL
KTPRVVGGEEASVDSWPWQVSIQYDKQHVCSSILDPHWVLTAAHCFRKHTDVFNWKVRAGSDKL
GSFPSLAVAKIIIIIEFNPMYPKDNDIALMKLQFPLTFSGTVRPICLPFFDEELTPATPLWIIIGWG
FTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPEGGVDTCCQDSSGGPLMYQS
DQWHVVGIVSWGYGCGGPSTPGVYTKVSAYLNWIYNVWKAEL

Transmembrane domain:

amino acids 32-53 (typeII)

FIGURE 113

GGCTGGACTGGAACCTCCTGGTCCCAAGTGATCCACCCGCCTCAGCCTCCCAAGGTGCTGTGATTA
 TAGGTGTAAGCCACCGTGTCTGGCCTCTGAACAACCTTTTTCAGCAACTAAAAAGCCACAGGAGT
 TGAAGTGTAGGATTCTGACTATGCTGTGGTGGCTAGTGCTCCTACTCCTACCTACATTAAATC
 TGTTTTTGTTCCTTGTAAGTACCTTTACCTTCCTAACACAGAGGATCTGTCACTGTGGCTCT
 GGCCCAAACCTGACCTTCACTCTGGAACGAGAACAGAGGTTTCTACCCACACCGTCCCTCGAAG
 CCGGGGACAGCCTCACCTTGCTGGCCTCTCGCTGGAGCAGTGCCCTCACCAACTGTCTCACGTCT
 GGAGGCACTGACTCGGGCAGTGCAGGTAGCTGAGCCTCTTGGTAGCTGCGGCTTCAAGGTGGGC
 CTTGCCCTGGCCGTAGAAGGGATTGACAAGCCCGAAGATTTTCATAGGCGATGGCTCCCACTGCC
 AGGCATCAGCCTTGCTGTAGTCAATCACTGCCCAGGAGCGGGCCGTGGACACCTGCTCA
 GAAGCAGTGGGTGAGACATCACGCTGCCCCGCCATCTAACCTTTTCATGTCTGCACATCACCTG
 ATCCATGGGCTAATCTGAAGTCTGTCCCAAGGAACCCAGAGCTTGAGTGAGCTGTGGCTCAGACC
 CAGAAGGGGTCTGCTTAGACCACCTGGTTTATGTGACAGGACTTGCACTTCTCCTGGAACATGAGG
 GAACGCCGGAGGAAAGCAAAGTGGCAGGGAAGGAAGTGTGCCAAATTATGGGTGAGAAAGATG
 GAGGTGTGGGTTATCACAAGGCATCGAGTCTCCTGCATTAGTGAGCATGTGGGGGAAGGGCTG
 CCGATGGCGCATGACACACTCGGGACTCACCTCTGGGGCCATCAGACAGCCGTTTCCGCCCCGAT
 CCACGTACCAGCTGCTGAAGGGCAACTGCAGGCCGATGCTCTCATCAGCCAGGCAGCAGCCAAA
 TCTGCGATCACACAGCCAGGGGCAGCCGTCTGGGAAGGAGCAAGCAAAGTGACCATTTCTCCTCCC
 CTCCTTCCCTCTGAGAGGCCCTCCTATGTCCCTACTAAAGCCACCAGCAAGACATAGCTGACAGG
 GGCTAATGGCTCAGTGTGGGCCAGGAGGTGAGCAAGGCCTGAGAGCTGATCAGAAGGGCCTGCT
 GTGCGAACACGGAAATGCCTCCAGTAAGCACAGGCTGCAAAATCCCCAGGCAAAGGACTGTGTGG
 CTCAATTTAAATCATGTTCTAGTAATTGGAGCTGTCCCCAAGACCAAAGGAGCTAGAGCTTGGTT
 CAAATGATCTCCAAGGGCCCTTATACCCAGGAGACTTTGATTTGAATTTGAAACCCCAATCCA
 AACCTAAGAACCAGGTGCATTAAGAATCAGTTATTGCCGGGTGTGGTGGCCTGTAATGCCAACAT
 TTTGGGAGGCCGAGGCGGTAGATCACCTGAGGTCAGGAGTTCAAGACCAGCTGGCCAACATGG
 TGAAACCCCTGTCTCTACTAAAAATACAAAAAACTAGCCAGGCATGGTGGTGTGTGCCTGTATC
 CCAGCTACTCGGGAGGCTGAGACAGGAGAATTACTTGAACCTGGGAGGTGAAGGAGGCTGAGACA
 GGAGAATCACTTCAGCCTGAGCAACACAGCGAGACTCTGTCTCAGAAAAAATAAAAAAGAATTA
 TGGTTATTTGTAA

[illegible]

Signal peptide:
amino acids 1-15

FIGURE 115

CAGCAGTGGTCTCTCAGTCCTCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACCATGGCAA
AGAATCCTCCAGAGAATTGTGAAGACTGTCACATTCTAAATGCAGAAGCTTTTAAATCCAAGAAA
ATATGTAAATCACTTAAGATTTGTGGACTGGTGTGGTATCCTGGCCCTAACTCTAATTGTCCT
GTTTGGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAGCCTATGACATGGAGCACACTT
TCTACAGCAATGGAGAGAAGAAGAAGATTACATGGAAATTGATCCTGTGACCAGAAGTGAATA
TTCAGAAGCGGAAATGGCACTGATGAAACATTGGAAGTGCACGACTTTAAAAACGGATACACTGG
CATCTACTTCGTGGGTCTTCAAAAATGTTTATCAAACTCAGATTAAAGTGATTCTGAATTTT
CTGAACCAGAAGAGGAAATAGATGAGAATGAAGAAATTACCACAACCTTCTTTGAACAGTCAGTG
ATTTGGGTCCCAGCAGAAAAGCCTATTGAAAACCGAGATTTCTTAAAAATTCAAAAATCTGGA
GATTGTGATAACGTGACCATGTATTGGATCAATCCCACTCTAATATCAGTTTCTGAGTTACAAG
ACTTTGAGGAGGAGGAGAAGATCTTCACTTTCCTGCCAACGAAAAAAGGGATTGAACAAAAT
GAACAGTGGGTGGTCCCTCAAGTGAAAGTAGAGAAGACCCGTCACGCCAGACAAGCAAGTGAGGA
AGAACTTCCAATAAATGACTATACTGAAAATGGAATAGAATTTGATCCCATGCTGGATGAGAGAG
GTTATTGTTGTATTTACTGCCGTCGAGGCAACCGCTATTGCCGCCGCTGTGAACCTTTACTA
GGCTACTACCCATATCCATACTGCTACCAAGGAGGACGAGTCATCTGTCGTGTCATCATGCCTTG
TAACTGGTGGGTGGCCCGCATGCTGGGGAGGGTCTAATAGGAGGTTTGAGCTCAAATGCTTAAAC
TGCTGGCAACATATAATAAATGCATGCTATTCAATGAATTTCTGCCTATGAGGCATCTGGCCCT
GGTAGCCAGCTCTCCAGAATTACTGTAGGTAATTCCTCTCTCATGTTCTAATAAACTTCTACA
TTATCACCAAAAAAAAAAAAAAAAAAAAA

FIGURE 116

MAKNPPENCEDCHILNAEAFKSKKICKSLKICGLVFGILALTLIVLFWGSKHFWPEVPKKAYDME
HTFYSNGEKKKIYMEIDPVTRTEIFRSGNGTDETLVHDFKNGYTGIIYFVGLQKCFIKTQIKVIP
EFSEPEEEIDENEEITTTFFEQSVIWPVPAEKPIENRDFLKNKILEICDNVTMYWINPTLISVSE
LQDFEEEGEDLHFPANEKKGIEQNEQWVVPQVKVEKTRHARQASEEELPINDYTENGIEFDPMLD
ERGYCCIIYCRNGNRYCRRVCEPLLGYYPYPYCYQGGRVICRVIMPCNWWVARMLGRV

Important features of the protein:

Signal peptide:

amino acids 1-40

Transmembrane domain:

amino acids 25-47 (type II)

N-glycosylation sites.

amino acids 94-97, 180-183

Glycosaminoglycan attachment sites.

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-242

N-myristoylation sites.

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

Microbodies C-terminal targeting signal.

amino acids 315-317

Cytochrome c family heme-binding site signature.

amino acids 9-14

[illegible]

MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVVFQYEGLRSCVRQSSGFTECRP
YFTILGLPAMLAQAVRALMIVGIVLGAIGLLVSI FALKCIRIGSMEDSAKANMTLTSGIMFIVSGL
CAIAGVSVFANMLVTNFWMSTANMYTGMGMVQTVQTRYTFGAALFVGWVAGGLTLIGGVMMCIA
CRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKIYDGGARTEDEVQSYPSKHDY
V

Signal peptide:

amino acids 1-23

Transmembrane domains:

amino acids 81-100, 121-141, 173-194

FIGURE 119

GGAAAACTGTTCTCTTCTGTGGCACAGAGAACCCTGCTTCAAAGCAGAAGTAGCAGTTCCGGAGTCC
 AGCTGGCTAAAACTCATCCCAGAGGATAATGGCAACCCATGCCTTAGAAATCGCTGGGCTGTTTCTTG
 GTGGTGTGGAATGGTGGGCACAGTGGCTGTCACTGTGCCTCAGTGGAGAGTGTCCGCCCTTCATT
 GAAAAACAACATCGTGGTTTTTGAAAACTTCTGGGAAGGACTGTGGATGAATTGCGTGAGGCAGGCTAA
 CATCAGGATGCAGTGCAAAATCTATGATTCCCTGCTGGCTCTTTCTCCGGACCTACAGGCAGCCAGAG
 GACTGATGTGTGCTGCTTCCGTGATGTCTTCTTGGCTTTCATGATGGCCATCCCTTGGCATGAAATGC
 ACCAGGTGCACGGGGGACAATGAGAAGGTGAAGGCTCACATTCTGCTGACGGCTGGAATCATCTTCAT
 CATCACGGGCATGGTGGTGTCTATCCCTGTGAGCTGGGTGCCAATGCCATCATCAGAGATTTCTATA
 ACTCAATAGTGAATGTTGCCCAAAACGTGAGCTTGGAGAAGCTCTCTACTTAGGATGGACCACGGCA
 CTGGTGTGATTGTTGGAGGAGCTCTGTTCTGCTGCGTTTTTTGTTGCAACGAAAAGAGCAGTAGCTA
 CAGATACTCGATACCTTCCCATCGCACAAACCAAAAAAGTTATCACACCGGAAAGAAGTCACCGAGCG
 TCTACTCCAGAAGTCAGTATGTGTAGTTGTGTATGTTTTTTAACTTTACTATAAAGCCATGCAAATG
 ACAAAAATCTATATTACTTTCTCAAAATGGACCCCAAGAACTTTGATTTACTGTTCTTAATGCGCT
 AATCTTAATTACAGGAAGTGTGCATCAGCTATTTATGATTCTATAAGCTATTTACAGCAGAATGAGATA
 TTAACCCCAATGCTTTTGATTGTTCTAGAAAGTATAGTAATTTGTTTTCTAAGGTGGTTCAAGCATCTA
 CTCTTTTTTATCATTTACTTCAAATGACATTGCTAAAGACTGCATTATTTTACTACTGTAATTTCTCC
 ACGACATAGCATTATGTACATAGATGAGTGTAAACATTTATATCTCACATAGAGACATGCTTATATGGT
 TTTATTTAAATGAAATGCCAGTCCATTACACTGAATAAATAGAACTCAACTATTGCTTTTCAGGGAA
 ATCATGGATAGGGTTGAAGAAGGTTACTATTAATTGTTTAAAAACAGCTTAGGGATTAATGTCTCCCA
 TTTATAATGAAGATTAATAATGAAGGCTTTAATCAGCATTGTAAAGGAAATTGAATGGCTTTCTGATAT
 GCTGTTTTTTAGCCTAGGAGTTAGAAATCCTAACTTCTTTATCCTCTTCTCCCAGAGGCTTTTTTTTT
 CTTGTGTATTAAATTAACATTTTTAAAAACGCAGATATTTTGTCAAGGGGCTTTGCATTCAAACGCTT
 TTCCAGGGCTATACTCAGAAGAAAGATAAAAGTGTGATCTAAGAAAAAGTGATGGTTTTAGGAAAGTG
 AAAATATTTTTGTTTTGTATTTGAAGAAGATGATGCATTTTGACAAGAAATCATATATGTATGGAT
 ATATTTTAATAAGTATTTGAGTACAGACTTTGAGGTTTCATCAATATAAATAAAGAGCAGAAAAATA
 TGTCTTGGTTTTTCATTTGCTTACCAAAAAACAACAACAAAAAAGTTGTCCTTTGAGAACTTCACCT
 GCTCCTATGTGGGTACCTGAGTCAAAATTGTCATTTTTGTTCTGTGAAAAATAAATTTCTCTTGTA
 CCATTTCTGTTTAGTTTTACTAAAAATCTGTAAATACTGTATTTTTCTGTTTATCCAAATTTGATGAA
 ACTGACAATCCAATTTGAAAGTTTGTGTGCGAGCTGTCTAGCTTAAATGAATGTGTTCTATTTGCTT
 TATACATTTATATTAATAAATTTGTACATTTTCTAATT

FIGURE 120

MATHALEIAGLFLGGVGMVGTVAVTVMPPQWRVSAFIENNIVVFENFW EGLWMNCVRQANIRMQCK
 IYDSLLALSPDLQAARGLMCAASVMSFLAFMAILGMKCTRCTGDNEKVKAHILLTAGIIFIITG
 MVVLIPVSWVANAIIRDFYNSIVNVAQKRELGEALYLGWTTALVLIVGGALFCCVFCNEKSSSY
 RYSIPSHRTTQKSYHTGKKSPSVYSRSQYV

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 82-101, 118-145, 164-188

FIGURE 121

GGAGAGAGGCGCGGGTGAAAGGCGCATTGATGCAGCCTGCGGGCGGCCTCGGAGCGGGCGGAG
CCAGACGCTGACCACGTTCTCTCGGTCTCCTCCGCCTCCAGCTCCGCGCTGCCCGGCAGCC
GGGAGCCATGGACCCAGGGCCCCGCGCCTCCCCGAGCGGCTCCGCGGCCTCCTGCTGCTCC
TGCTGCTGCAGCTGCCCGCGCCGTCGAGCGCCTCTGAGATCCCCAAGGGGAAGCAAAAGGCGCAG
CTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGCTTACAAGGGCCAGCAGGAGTGCC
TGGTCGAGACGGGAGCCCTGGGGCCAATGTTATTCCGGGTACACCTGGGATCCCAGGTCGGGATG
GATTCAAAGGAGAAAAGGGGAATGTCTGAGGGAAAGCTTTGAGGAGTCCTGGACACCCAACCTAC
AAGCAGTGTTTCATGGAGTTCATTGAATTATGGCATAGATCTTGGGAAAATTGCGGAGTGATACATT
TACAAAGATGCGTTCAAATAGTGCTCTAAGAGTTTTGTTTCAGTGGCTCACTTCGGCTAAAATGCA
GAAATGCATGCTGTCAGCGTTGGTATTTACATTCAATGGAGCTGAATGTTTCAGGACCTCTTCCC
ATTGAAGCTATAATTTATTTGGACCAAGGAAGCCCTGAAATGAATTCAACAATTAATATTCATCG
CACTTCTTCTGTGGAAGGACTTTGTGAAGGAATTGGTGCTGGATTAGTGGATGTTGCTATCTGGG
TTGGCACTTGTTTCAGATTACCCAAAAGGAGATGCTTCTACTGGATGGAATTCAGTTTCTCGCATC
ATTATTGAAGAACTACCAAAATTAAATGCTTTAATTTTCATTGCTACCTCTTTTTTTATTATGCC
TTGGAATGGTTCACTTAAATGACATTTTAAATAAGTTTATGTATACATCTGAATGAAAAGCAAAG
CTAAATATGTTTACAGACCAAAGTGATTTACACTGTTTTTAAATCTAGCATTATTCATTTTG
CTTCAATCAAAGTGTTTCAATATTTTTTTTAGTTGGTTAGAATACTTCTTCATAGTCACATT
CTCTCAACCTATAATTTGGAATATTGTTGTGGTCTTTTGTTTTTTCTCTTAGTATAGCATTTTTTA
AAAAAATATAAAAGCTACCAATCTTTGTACAATTTGTAAATGTTAAGAATTTTTTTTATATCTGT
TAAATAAAAATTATTTCCAACA

FIGURE 122

MRPQGPAASPQRLRGLLLLLLLQLPAPSSASEIPKGKQKAQLRQREVVDLYNGMCLQGPAGVPGR
 DGSPGANVIPGTPGIPGRDGFKEGKECLRESFEESWTPNYKQCSWSSLNYGIDLKGIAECTFTK
 MRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECSGPLPIEAIYLDQGSPEMNSTINIHRTS
 SVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRIIEELPK

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

FIGURE 123

GCTGAGCGTGTGCGCGGTACGGGGCTCTCCTGCCTTCTGGGCTCCAACGCAGCTCTGTGGCTGAA
 CTGGGTGCTCATCACGGGAACGTGCTGGGCTATGGAATACAGATGTGGCAGCTCAGGTAGCCCCAA
 ATTGCCTGGAAGAATACATCATGTTTTTCGATAAGAAGAAATTGTAGGATCCAGTTTTTTTTTTA
 ACCGCCCCCTCCCCACCCCCCAAAAAAAGCTGTAAAGATGCAAAAACGTAATATCCATGAAGATCC
 TATTACCTAGGAAGATTTTGATGTTTTGCTGCGAATGCGGTGTTGGGATTTATTTGTTCTTGGAG
 TGTCTGCGTGGCTGGCAAAGAATAATGTTCCAAAATCGGTCCATCTCCCAAGGGGTCCAATTTT
 TCTTCTGGGTGTCAGCGAGCCCTGACTCACTACAGTGCAGCTGACAGGGGCTGTCTATGCAACTG
 GCCCTAAGCCAAAGCAAAAGACCTAAGGACGACCTTTGAACAATACAAAGGATGGGTTTCAATG
 TAATTAGGCTACTGAGCGGATCAGCTGTAGCACTGGTTATAGCCCCACTGTCTTACTGACAATG
 CTTTCTTCTGCCGAACGAGGATGCCCTAAGGGCTGTAGGTGTGAAGGCAAAATGGTATATTGTGA
 ATCTCAGAAATTACAGGAGATACCCCTCAAGTATATCTGCTGGTTGCTTAGGTTTGTCCCTTCGCT
 ATAACAGCCTTCAAAAACCTTAAGTATAATCAATTTAAAGGGCTCAACCAGCTCACCTGGCTATAC
 CTTGACCATAACCATATCAGCAATATTGACGAAAATGCTTTTAAATGGAATACGCAGACTCAAAGA
 GCTGATTCTTAGTTCCAATAGAATCTCCTATTTCTTAAACAATACCTTCAGACCTGTGACAAATT
 TACGGAACTTGGATCTGTCTATAATCAGCTGCATTCTCTGGGATCTGAACAGTTTCGGGGCTTG
 CGGAAGCTGCTGAGTTTACATTACGGTCTAACTCCCTGAGAACCATCCCTGTGCGAATATTCCA
 AGACTGCCGCAACCTGGAACCTTTGGACCTGGGATATAACCGGATCCGAAGTTTAGCCAGGAATG
 TCTTTGCTGGCATGATCAGACTCAAAGAACTTCACCTGGAGCACAATCAATTTTCCAAGCTCAAC
 CTGGCCCTTTTTTCCAAGGTTGGTCAGCCTTCAGAACCTTTACTTGCAGTGAATAAAATCAGTGT
 CATAGGACAGACCATGTCTGGACCTGGAGCTCCTTACAAAGGCTTGATTTATCAGGCAATGAGA
 TCGAAGCTTTTCAGTGGACCCAGTGTTTTCCAGTGTGTCCCGAATCTGCAGCGCCTCAACCTGGAT
 TCCAACAAGCTCACATTTATTGGTCAAGAGATTTTGGATTCTTGGATATCCCTCAATGACATCAG
 TCTTGCTGGGAATATATGGGAATGCAGCAGAAATATTTGCTCCCTTGTAACCTGGCTGAAAAGTT
 TTAAAGGTCTAAGGGAGAATACAATTATCTGTGCCAGTCCCAAAGAGCTGCAAGGAGTAAATGTG
 ATCGATGCAGTGAAGAACTACAGCATCTGTGGCAAAAGTACTACAGAGAGGTTTGATCTGGCCAG
 GGCTCTCCCAAAGCCGACGTTTAAAGCCAAAGCTCCCCAGGCCGAAGCATGAGAGCAAACCCCTT
 TGCCCCCGACGGTGGGAGCCACAGAGCCCGGCCAGAGACCGATGCTGACGCCGAGCACATCTCT
 TTCCATAAAATCATCGCGGGCAGCGTGGCGCTTTTCTGTCCGTGCTCGTCATCCTGCTGGTTAT
 CTACGTGTCATGGAAGCGGTACCTTGGGAGCATGAAGCAGCTGCAGCAGCGCTCCCTCATGCGAA
 GGCACAGGAAAAAGAAAAGACAGTCCCTAAAGCAAATGACTCCAGCACCCAGGAATTTTATGTA
 GATTATAAACCACCAACACGGAGACCAGCGAGATGCTGCTGAATGGGACGGGACCCCTGCACCTA
 TAACAAATCGGGCTCCAGGGAGTGTGAGGTAAGAACCATTTGTGATAAAAAGAGCTCTTAAAGCT
 GGGAAATAAGTGGTGTCTTATTGAACTCTGGTGACTATCAAGGGAACGCGATGCCCCCTCCCC
 TTCCCTCTCCCTCTCACTTTGGTGGAAGATCCTTCCTTGTCCGTTTTAGTGCATTATATAACT
 GGTCAATTTTCTCTCATACATAATCAACCCATTGAAATTTAAATACCACAATCAATGTGAAGCTT
 GAATCCGGTTTAAATATAATACCTATTGTATAAGACCCCTTACTGATTCCATTAATGTGCGATTT
 GTTTTAAGATAAAACTTCTTTCATAGGTAAAAA

FIGURE 124

MGFNVIRLLSGSAVALVIAPTVLLTMLSSAERGCPKGCRCEGKMVYCESQKLQEIPSSISAGCLG
LSLRYNLSLQKLKYNQFKGLNQLTWLYLDHNHISNIDENAFNGIRRLKELILSSNRISYFLNNTFR
PVTNLRNLDSLQNQLHSLGSEQFRGLRKLLSLHLRSNSLRTIPVRIQDCRNLELLDLGYNRIRS
LARNVFAGMIRLKEHLEHNQFSKLNLFRLVSLQNLQWNKISVIGQTMSTWSSLRDL
SGNEIEAFSGPSVFQCVPNLQRLNLDNSNKLTFIGQEILDSWISLNDISLAGNIWECSRNICSLVN
WLKSFKGLRENTIICASPKELQGVNVIDAVKNYSICGKSTTERFDLARALPKPTFKPKLPRPKHE
SKPPLPPTVGATEPGPETDADAHEHISFHKIIAGSVALFLSVLVILLVIYVSWKRYPASMKQLQQR
SLMRHRHKKRQSLKQMTPTSTQEFYVDYKPTNTETSEMLLNGTGPCTYNKSGSRECEV

Important features of the protein:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 420-442

N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

Tyrosine kinase phosphorylation site.

amino acids 136-142

N-myristoylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

[illegible]

App_ID=10063521

FIGURE 126

MSVPEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMGEAALARLGDGARES
 APYRGMVRTALGIIIEEGFLKLWQGVTPAIYRHVVYSGGRMVTYEHLEVVFGKSEDEHYPLWKS
 VIGGMMAGVIGQFLANPTDLVKVQMOMEGKRKLEGKPLRFRGVHHAFAKILAEGGIRGLWAGWVP
 NIQRAALVNMGDLTTYDTVKHYLVLNTPLEDNIMTHGLSSLCSGLVASILGTPADVIKSRIMNQP
 RDKQGRGLLYKSSTDCLIQAVQGEGFMSLYKGFLPSWLRMTPWSMVFWLTYEKIREMSGVSPF

Transmembrane domains:

amino acids 25-38, 130-147, 233-248

FIGURE 127

CGCGGATCGGACCCAAGCAGGTCGGCGGGCGGCGGAGAGCGGCCGGGCGTCAGCTCCTCGAC
 CCCCCTGTCGGGCTAGTCCAGCGAGGCGGACGGGCGGCGTGGGCCCATGGCCAGGCCCGGCATGG
 AGCGGTGGCGCGACCGGCTGGCGCTGGTGACGGGGGCCTCGGGGGGCATCGGCGCGGCCGTGGCC
 CGGGCCCTGGTCCAGCAGGGACTGAAGGTGGTGGGCTGCGCCCGCACTGTGGGCAACATCGAGGA
 GCTGGCTGCTGAATGTAAGAGTGCAGGCTACCCCGGGACTTTGATCCCCACAGATGTGACCTAT
 CAAATGAAGAGGACATCCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGTGTAGACATC
 TGCATCAACAATGCTGGCTTGGCCCGGCTGACACCCTGCTCTCAGGCAGCACCAGTGGTTGGAA
 GGACATGTTCAATGTGAACGTGCTGGCCCTCAGCATCTGCACACGGGAAGCCTACCAGTCCATGA
 AGGAGCGGAATGTGGACGATGGGCACATCATTAACATCAATAGCATGTCTGGCCACCGAGTGTTA
 CCCCTGTCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCGTCACCTGCGCTGACAGAGGGACT
 GAGGCAAGAGCTTCGGGAGGCCCAGACCCACATCCGAGCCACGTGCATCTCTCCAGGTGTGGTGG
 AGACACAATTGCGCTTCAAACCTCCACGACAAGGACCCTGAGAAGGCAGCTGCCACCTATGAGCAA
 ATGAAGTGCTCAAACCCGAGGATGTGGCCGAGGCTGTTATCTACGTCCCTCAGCACCCCCGCACA
 CATCCAGATTGGAGACATCCAGATGAGGCCCACGGAGCAGGTGACCTAGTGACTGTGGGAGCTCC
 TCCTTCCCTCCCCACCCCTTCATGGCTTGCCTCCTGCCTCTGGATTTTAGGTGTTGATTTCTGGAT
 CACGGGATACCACTTCCTGTCCACACCCCGACCAGGGGCTAGAAAATTTGTTTGAGATTTTATA
 TCATCTTGTCAAATTGCTTCAGTTGTAAATGTGAAAAATGGGCTGGGGAAAGGAGTGGTGTCCC
 TAATTGTTTTACTTGTTAACTTGTTCTTGTGCCCTGGGCACTTGGCCTTTGTCTGCTCTCAGTG
 TCTTCCCTTTGACATGGGAAAGGAGTTGTGGCCAAATCCCCATCTTCTTGACCTCAACGTCTG
 TGGCTCAGGGCTGGGGTGGCAGAGGGAGGCCTTCACCTTATATCTGTGTTGTTATCCAGGGCTCC
 AGACTTCCTCCTCTGCCTGCCCCACTGCACCCTCTCCCCCTTATCTATCTCCTTCTCGGCTCCCC
 AGCCCAGTCTTGGCTTCTTGTCCCCTCCTGGGGTCATCCCTCCACTCTGACTCTGACTATGGCAG
 CAGAACACCAGGGCCTGGCCCAGTGGATTTTCATGGTGATCATTAAGAAAGAAAAATCGCAACCAA
 AAAAAAAAAA

FIGURE 128

MARPGMERWRDRLALVTGASGGIGA A VARALVQQGLKVVG CARTVGNIEELAAECKSAGYPGTLI
PYRCDLSNEEDILSMFSAIRSQHSGVDICINNAGLARPD TLLSGSTSGWKDMFN VNVLALSICTR
EAYQSMKERNVDDGHIININMSGHRVLP LSVTHFY SATKYAVTALTEGLRQELREAQT HIRATC
ISPGVVETQFAFKLHDKDPEKAAATYEQMKCLK PEDVAEAVIYVLSTPAHIQIGDIQMRPTEQVT

Important features of the protein:

Signal peptide:

amino acids 1-17

N-myristoylation sites.

amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115,
199-205

Short-chain alcohol dehydrogenase.

amino acids 30-42, 104-114

FIGURE 129

AACTTCTACATGGGCCTCCTGCTGCTGGTGGCTCTTCTCAGCCTCCTGCCGGTGGCCTACACCAT
CATGTCCCTCCCACCCTCCTTTGACTGCGGGCCGTTCAAGGTGCAGAGTCTCAGTTGCCCGGGAGC
ACCTCCCCTCCCAGGCAGTCTGCTCAGAGGGCCTCGGCCAGAAATCCAGTTCTGGTTTCATGC
CAGCCTGTAAAAGGCCATGGAACCTTTGGGTGAATCACCGATGCCATTTAAGAGGGTTTTCTGCCA
GGATGGAAATGTTAGGTCGTTCTGTGTCTGCGCTGTTCAATTCAGTAGCCACCAGCCACCTGTGG
CCGTTGAGTGCTTGAAATTAGGAACTGAGAAAATTAATTTCTCATGTATTTTTCTCATTTATTTA
TTAATTTTTAACTGATAGTTGTACATATTTGGGGGTACATGTGATATTTGGATACATGTATACAA
TATATAATGATCAAATCAGGGTAACTGGGATATCCATCACATCAAACATTTATTTTTTATTCTTT
TTAGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGTGGTGCCATCTCAGCTTACTGCAAC
CTCTGCCTGCCAGGTTCAAGCGATTCTCATGCCTCCACCTCCCAAGTAGCTGGGACTACAGGCAT
GCACCACAATGCCCAACTAATTTTTGTATTTTTAGTAGAGACGGGGTTTTGCCATGTTGCCCAGG
CTGGCCTTGAACTCCTGGCCTCAAACAATCCACTTGCCTCGGCCTCCCAAAGTGTTATGATTACA
GGCGTGAGCCACCGTGCCTGGCCTAAACATTTATCTTTTCTTTGTGTTGGGAACCTTGAAATTAT
ACAATGAATTATTTGTTAACTGTCATCTCCCTGCTGTGCTATGGAACACTGGGACTTCTTCCCTCT
ATCTAACTGTATATTTGTACCAGTTAACCAACCGTACTTCATCCCCACTCCTCTCTATCCTTCCC
AACCTCTGATCACCTCATTCTACTCTCTACCTCCATGAGATCCACTTTTTTAGCTCCCATGTG
AGTAAGAAAATGCAATATTTGTCTTTCTGTGCCTGGCTTATTTCACTTAACATAATGACTTCCTG
TTCCATCCATGTTGCTGCAAATGACAGGATTCGTTCTTAATTTCAATTAAAAATAACCACACATG
GCAAAAA

$$C_{\text{eff}} = \frac{\sum_{j=1}^n C_j}{n}$$

Important features of the protein:

amino acids 1-18

amino acids 86-92

amino acids 68-79

FIGURE 131

TTCTGAAGTAACGGAAGCTACCTTGTATAAAGACCTCAACACTGCTGACCATGATCAGCGCAGCCTGGAGC
ATCTTCCTCATCGGGACTAAAATTGGGCTGTTCCCTCAAGTAGCACCTCTATCAGTTATGGCTAAATCCTG
TCCATCTGTGTGCTGCGATGCGGGTTTCATTTACTGTAATGATCGCTTTCTGACATCCATTCCAACAG
GAATACCAGAGGATGCTACAACCTCTACCTTCAGAACAACCAAATAAATAATGCTGGGATTCCCTCAGAT
TTGAAAAACTTGCTGAAAGTAGAAAGAATATACCTATACCACAACAGTTTAGATGAATTCCTACCAACCT
CCCAAAGTATGTAAAAGAGTTACATTTGCAAGAAAATAACATAAGGACTATCACTTATGATTCACCTTTCAA
AAATTCCTATCTGGAAGAATTACATTTAGATGACAACTCTGTCTCTGCAGTTAGCATAGAAGAGGGAGCA
TTCCGAGACAGCAACTATCTCCGACTGCTTTTCCTGTCCCGTAATCACCTTAGCACAAATCCCTGGGGTTT
GCCCAGGACTATAGAAGAACTACGCTTGGATGATAATCGCATATCCACTATTTTCATCACCATCTCTTCAAG
GTCTCACTAGTCTAAAACGCCCTGGTCTAGATGGAAACCTGTTGAACAATCATGGTTTAGGTGACAAAGTT
TTCTTCAACCTAGTTAATTTGACAGAGCTGTCCCTGGTGCGGAATTCCCTGACTGCTGCACCAGTAAACCT
TCCAGGCACAAACCTGAGGAAGCTTTATCTTCAAGATAACCACATCAATCGGGTGCCCCCAATGCTTTTTT
CTTATCTAAGGCAGCTCTATCGACTGGATATGTCCAATAATAACCTAAGTAATTTACCTCAGGGTATCTTT
GATGATTTGGACAATATAACACAACCTGATTCTTCGCAACAATCCCTGGTATTGCGGGTGCAAGATGAAATG
GGTACGTGACTGGTTACAATCACTACCTGTGAAGGTCAACGTGCGTGGGCTCATGTGCCAAGCCCCAGAAA
AGGTTCTGGGGATGGCTATTAAGGATCTCAATGCAGAACTGTTTGATTGTAAGGACAGTGGGATTGTAAGC
ACCATTAGATAACCACTGCAATACCCAACACAGTGTATCCTGCCAAGGACAGTGGCCAGCTCCAGTGAC
CAAACAGCCAGATATTAAGAACCCCAAGCTCACTAAGGATCAACAACCACAGGGAGTCCCTCAAGAAAAA
CAATTACAATTACTGTGAAGTCTGTACCTCTGATACCATTATATCTCTTGGAACTTGCTCTACCTATG
ACTGCTTTGAGACTCAGCTGGCTTAAACTGGGCCATAGCCCGGCATTTGGATCTATAACAGAAACAATTGT
AACAGGGGAACGCAGTGAGTACTTGGTCACAGCCCTGGAGCCTGATTACCCCTATAAAGTATGCATGGTTC
CCATGGAAACCAGCAACCTCTACCTATTTGATGAAACTCCTGTTTGATTGAGACTGAACTGCACCCCTT
CGAATGTACAACCCTACAACCACCTCAATCGAGAGCAAGAGAAAGAACCTTACAAAAACCCCAATTTACC
TTTGGCTGCCATCATTTGGTGGGCTGTGGCCCTGGTTACCATTGCCCTTCTTGCTTTAGTGTGTGGTATG
TTCATAGGAATGGATCGCTCTTCTCAAGGAACCTGTGCATATAGCAAAGGGAGGAGAAGAAAGGATGACTAT
GCAGAAGCTGGCACTAAGAAGGACAACCTCTATCCTGGAAATCAGGGAAACTTCTTTTCAGATGTTACCAAT
AAGCAATGAACCCATCTCGAAGGAGGAGTTTGTAAATACACACCATATTTCCCTCCTAATGGAATGAATCTGT
ACAAAAACAATCACAGTGAAGCAGTAGTAACCGAAGCTACAGAGACAGTGGTATTCCAGACTCAGATCAC
TCACACTCATGATGCTGAAGGACTCACAGCAGACTTGTGTTTTGGGTTTTTTAAACCTAAGGGAGGTGATG
GT

FIGURE 132

MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVCRC DAGFIYCND RFLTSIPTGIPEDATTLYL
QNNQINNAGIPSDLKNLLKVERIYLYHNSLDEFPTNLPKYVKELHLQENNIR TITYDSLSKIPYL
EELHDDNSVSAVSIEEGA FRDSNYLRLLFLSRNHLSTIPWGLPRTIEELRLDDNRISTISSPSL
QGLTSLKRLVLDGNLLNNHGLGDKVFFNLVNLTELSLVRNSLTAAPVNLPGTNLRKLYLQDNHIN
RVPPNAFSYLRQLYRLDMSNNNLSNLPQGI FDDL DNITQLILRN NPWYCGCKMKWVRDWLQSLPV
KVNVRGLMCQAPEKVRGMAIKDLNAELFDCKDSGIVSTIQITTAIPNTVYPAQGQWPAPVTKQPD
IKNPKLTKDQQTGSPSRKTITITVKSVTSDTIHISWKLALPMTALRLSWLKLGHSPA FGSITET
IVTGERSEYLVTALEPDSPYKVCMPMETS NLYLFD ETPVC IETETAPLRMYNP TTTTLNREQEKE
PYKNPNPLAAIIGGAVALVTIAL LALVCWYVHRNGSLFSRNCAYSKGRRRKDDYAEAGTKKDNS
ILEIRETSFQMLPISNEPISKEEFVIHTIFPPNGMNL YKNNHSESSSNRSYRDSGIPDS DHSHS

Important features of the protein:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 531-552

N-glycosylation sites.

amino acids 226-229, 282-285, 296-299, 555-558, 626-629, 633-636

Tyrosine kinase phosphorylation site.

amino acids 515-522

N-myristoylation sites.

amino acids 12-17, 172-177, 208-213, 359-364, 534-539, 556-561,
640-645

Amidation site.

amino acids 567-570

Leucine zipper pattern.

amino acids 159-180

Phospholipase A2 aspartic acid active site.

amino acids 34-44

FIGURE 133

CCGTCATCCCCCTGCAGCCACCCTTCCCAGAGTCCTTTGCCCAGGCCACCCAGGCTTCTTGCCA
 GCCCTGCCGGGCCACTTGTCTTCATGTCTGCCAGGGGGAGGTGGGAAGGAGGTGGGAGGAGGGCG
 TGCAGAGGCAGTCTGGGCTTGGCCAGAGCTCAGGGTGCTGAGCGTGTGACCAGCAGTGAGCAGAG
 GCCGGCCATGGCCAGCCTGGGGCTGCTGCTCCTGCTCTTACTGACAGCACTGCCACCGCTGTGGT
 CCTCCTCACTGCCTGGGCTGGACACTGCTGAAAGTAAAGCCACCATTGCAGACCTGATCCTGTCT
 GCGCTGGAGAGAGCCACCGTCTTCCTAGAACAGAGGCTGCCTGAAATCAACCTGGATGGCATGGT
 GGGGGTCCGAGTGCTGGAAGAGCAGCTAAAAAGTGTCGGGAGAAGTGGGCCAGGAGCCCCCTGC
 TGCAGCCGCTGAGCCTGCGCGTGGGGATGCTGGGGGAGAAGCTGGAGGCTGCCATCCAGAGATCC
 CTCCACTACCTCAAGCTGAGTGATCCCAAGTACCTAAGAGAGTTCCAGCTGACCCCTCCAGCCCGG
 GTTTTGGGAAGCTCCACATGCCTGGATCCACACTGATGCCTCCTTGGTGTACCCACGTTCCGGGC
 CCCAGGACTCATTCTCAGAGGAGAGAAGTGACGTGTGCCTGGTGCAGCTGCTGGGAACCGGGACG
 GACAGCAGCGAGCCCTGCGGCCTCTCAGACCTCTGCAGGAGCCTCATGACCAAGCCCGGCTGCTC
 AGGCTACTGCCTGTCCCACCAACTGCTCTTCTCCTCTGGGCCAGAATGAGGGGATGCACACAGG
 GACCACTCCAACAGAGCCAGGACTATATCAACCTCTTCTGCGCCAACATGATGGACTTGAACCGC
 AGAGCTGAGGCCATCGGATACGCCTACCCACCCGGACATCTTCATGGAAAACATCATGTTCTG
 TGGAATGGGCGGCTTCTCCGACTTCTACAAGCTCCGGTGGCTGGAGGCCATTCTCAGCTGGCAGA
 AACAGCAGGAAGGATGCTTCGGGGAGCCTGATGCTGAAGATGAAGAATTATCTAAAGCTATTCAA
 TATCAGCAGCATTTTTCGAGGAGAGTGAAGAGGCGAGAAAAACAATTTCCAGATTCTCGCTCTGT
 TGCTCAGGCTGGAGTACAGTGGCGCAATCTCGGCTCACTGCAACCTTTGCCTCCTGGGTTCAAGC
 AATTCTCTTGCTCATCCTCCCAGTAGCTGGGACTACAGGAGCGTGCCACCATACTGGCTAAT
 TTTTATATTTTTTTAGTAGAGACAGGGTTTCATCATGTTGCTCATGCTGGTCTCGAACTCCTGAT
 CTCAAGAGATCCGCCCACCTCAGGCTCCCAAAGTGTGGGATTATTAGGTGTGAGCCACCGTGTCTG
 GCTGAAAAGCACTTTCAAAGAGACTGTGTTGAATAAAGGGCCAAGGTTCTTGCCACCCAGCACTC
 ATGGGGGCTCTCTCCCTAGATGGCTGCTCCTCCACAACACAGCCACAGCAGTGGCAGCCCTGG
 GTGGCTTCTATACATCCTGGCAGAATACCCCCAGCAAACAGAGAGCCACACCCATCCACACCG
 CCACCACCAAGCAGCCGCTGAGACGGACGGTTCCATGCCAGCTGCCTGGAGGAGGAACAGACCCC
 TTTAGTCCTCATCCCTTAGATCCTGGAGGGCACGGATCACATCCTGGGAAGAAGGCATCTGGAGG
 ATAAGCAAAGCCACCCGACACCCAATCTTGAAGCCCTGAGTAGGCAGGGCCAGGGTAGGTGGG
 GGCCGGGAGGGACCCAGGTGTGAACGGATGAATAAAGTTCAACTGCAACTGAAAAAAAAAA

FIGURE 134

MSARGRWEGGRRACRGSGLLARAQGAERTVSSEQRPAMASLGLLLLLLLTALPPLWSSSLPGLD
 TAESKATIADLILSALERATVFLEQRLPEINLDGMVGVRVLEEQLKSVREKWAQEPLQLPLSLRV
 GMLGEKLEAAIQRSYLHLKSLDPKYLREFQLTLQPGFWKLPHAWIHTDASLVYPTFGPQDSFSEE
 RSDVCLVQLLGTGTDSEPCGLSDLCRSLMTKPGCSGYCLSHQLLFFLWARMRGCTQGPLQQSQD
 YINLFCANMMDLNRRAEAIGYAYPTRDIFMENIMFCGMGGFSDFYKLRWLEAILSQKQEGCFG
 EPDAEDEELSKAIQYQQHFSSRRVKRREKQFPDSRSVAQAGVQWRNLGSLQPLPPGFKQFSLILP
 SSWDYRSVPPYLANFYIFLVETGFHHVAHAGLELLISRDPPPTSGSQSVGL

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 39-56

Tyrosine kinase phosphorylation sites.

amino acids 149-156, 274-282

N-myristoylation sites.

amino acids 10-16, 20-26, 63-69, 208-214

Amidation site.

amino acids 10-14

Glycoprotein hormones beta chain signature 1.

amino acids 230-237

FIGURE 135

GGTCTGAGTGCAGAGCTGCTGTCATGGCGGCGCTCTGTGGGGCTTCTTTCCCGTCCCTGCTGCTG
CTGCTGCTATCGGGGGATGTCCAGAGCTCGGAGGTGCCCGGGGCTGCTGCTGAGGGATCGGGAGG
GAGTGGGGTCGGCATAGGAGATCGCTTCAAGATTGAGGGGCGTGCAAGTTGTTCCAGGGGTGAAGC
CTCAGGACTGGATCTCGGCGGCCCGAGTGCTGGTAGACGGAGAAGAGCACGTCGGTTTCCTTAAG
ACAGATGGGAGTTTTGTGGTTCATGATATACCTTCTGGATCTTATGTAGTGGAAGTTGTATCTCC
AGCTTACAGATTTGATCCCGTTCGAGTGGATATCACTTCGAAAGGAAAAATGAGAGCAAGATATG
TGAATTACATCAAAACATCAGAGGTTGTCAGACTGCCCTATCCTCTCCAAATGAAATCTTCAGGT
CCACCTTCTTACTTTATTAAAAGGGAATCGTGGGGCTGGACAGACTTTCTAATGAACCCAATGGT
TATGATGATGGTTCTTCCTTTATTGATATTTGTGCTTCTGCCTAAAGTGGTCAACACAAGTGATC
CTGACATGAGACGGGAAATGGAGCAGTCAATGAATATGCTGAATTCCAACCATGAGTTGCCTGAT
GTTTCTGAGTTCATGACAAGACTCTTCTCTTCAAAATCATCTGGCAAATCTAGCAGCGGCAGCAG
TAAACAGGCAAAAGTGGGGCTGGCAAAGGAGGTAGTCAGGCCGTCCAGAGCTGGCATTTGCAC
AAACACGGCAACACTGGGTGGCATCCAAGTCTTGAAAACCGTGTGAAGCAACTACTATAAACTT
GAGTCATCCCGACGTTGATCTCTTACAACGTGTATGTT
AACTTTTTAGCACATGTTTTGTACTTGGTACACGAGAAAACCCAGCTTTCATCTTTTGTCTGTAT
GAGGTCAATATTGATGTCACCTGAATTAATTACAGTGTCTATAGAAAATGCCATTAATAAATTAT
ATGAACTACTATACATTATGTATATTAATTAACATCTTAATCCAGAAATCAAAAAAAAAAAAAA
AAAAAAAAAAAAAA

FIGURE 136

MAAALWGFFPVLLLLLLSGDVQSSEVPGAAAEGSGSGVGIGDRFKIEGRAVVPGVKPQDWISAA
 RVLVDGEEHVGFLLKTDGSFVVDIPSGSYVVEVVSPAYRFDPVRVDITSKGKMRARYVNYIKTSE
 VVRLPYPLQMKSSGPPSYFIKRESWGWTDFLMNPMVMMVLPPLIFVLLPKVVNTSDPDMRREME
 QSMNMLNSNHELDPVSEFMTRLFSSKSSGKSSSGSSKTGKSGAGKRR

Important features of the protein:

Signal sequence:

amino acids 1-23

Transmembrane domain:

amino acids 161-182

N-glycosylation site.

amino acids 184-187

Glycosaminoglycan attachment sites.

amino acids 37-40, 236-239

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 151-154

N-myristoylation sites.

amino acids 33-38, 36-41, 38-44, 229-234

Amidation site.

amino acids 238-241

ATP/GTP-binding site motif A (P-loop).

amino acids 229-236

FIGURE 137

GATGGCGCAGCCACAGCTTCTGTGAGATTTCGATTTCTCCCCAGTTCCTGTGGGTCTGAGGGGA
 CCAGAAGGGTGAGCTACGTTGGCTTTCTGGAAGGGGAGGCTATATGCGTCAATTCCCCAAAACAA
 GTTTTGACATTTCCCTGAAATGTCATTCTCTATCTATTCACTGCAAGTGCCTGCTGTTCCAGGC
 CTTACCTGCTGGGCACTAACGGCGGAGCCAGGATGGGGACAGAATAAAGGAGCCACGACCTGTGC
 CACCAACTCGCACTCAGACTCTGAACTCAGACCTGAAATCTTCTCTTACGGGAGGCTTGGCAGT
 TTTTCTTACTCCTGTGGTCTCCAGATTTCAGGCCTAAGATGAAAGCCTCTAGTCTTGCCTTCAGC
 CTTCTCTCTGCTGCGTTTTATCTCCTATGGACTCCTTCCACTGGACTGAAGACACTCAATTTGGG
 AAGCTGTGTGATCGCCACAAACCTTCAGGAAATACGAAATGGATTTTCTGAGATACGGGGCAGTG
 TGCAAGCCAAAGATGGAACATTGACATCAGAATCTTAAGGAGGACTGAGTCTTTGCAAGACACA
 AAGCCTGCGAATCGATGCTGCCTCCTGCGCCATTTGCTAAGACTCTATCTGGACAGGGTATTTAA
 AAATACCAGACCCCTGACCATTAATCTCTCCGGAAGATCAGCAGCCTCGCCAATTCCTTTCTTA
 CCATCAAGAAGGACCTCCGGCTCTCTCATGCCCACATGACATGCCATTGTGGGGAGGAAGCAATG
 AAGAAATACAGCCAGATTCTGAGTCACTTTGAAAAGCTGGAACCTCAGGCAGCAGTTGTGAAGGC
 TTTGGGGGAAGTAGACATTCTTCTGCAATGGATGGAGGAGACAGAATAGGAGGAAAGTGATGCTG
 CTGCTAAGAATATTCGAGGTCAAGAGCTCCAGTCTTCAATACCTGCAGAGGAGGCATGACCCCAA
 ACCACCATCTCTTTACTGTACTAGTCTTGTGCTGGTCACAGTGTATCTTATTTATGCATTACTTG
 CTTCTTGCATGATTGTCTTTATGCATCCCCAATCTTAATTGAGACCATACTTGTATAAGATTTT
 TGTAATATCTTTCTGCTATTGGATATATTTATTAGTTAATATATTTATTTATTTTTTGTATTTA
 ATGTATTTATTTTTTTTACTTGGACATGAACTTTAAAAAATTACAGATTATATTTATAACCTG
 ACTAGAGCAGGTGATGTATTTTTATACAGTAAAAAAAAAACCTTGTAATTTCTAGAAGAGTGG
 CTAGGGGGGTATTCATTTGTATTCAACTAAGGACATATTTACTCATGCTGATGCTCTGTGAGAT
 ATTTGAAATTGAACCAATGACTACTTAGGATGGGTGTGGAATAAGTTTGTGTGGAATTGCAC
 ATCTACCTTACAATTACTGACCATCCCCAGTAGACTCCCCAGTCCCATAATTGTGTATCTCCAG
 CCAGGAATCCTACACGCCAGCATGTATTTCTACAAATAAAGTTTCTTTGCATACCAAAAAAAA
 AAAAAAAAAA

FIGURE 138

MRQFPKTSFDISPEMSFSIYSLQVPAVPGLTCWALTAEPGWGQNKGATTCATNSHSDSELRPEIF
 SSREAWQFFLLLWSPDFRPKMKASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNLQEIRNG
 FSEIRGSVQAKDGNIDIRILRRTESLQDTKPANRCCLLRHLLRLYLDRVFKNYQTPDHYTLRKIS
 SLANSFLTIKKDLRLSHAMTCHCGEAMKKYSQILSHFEKLEPQAAVVKALGELDILLQWMEET
 E

Important features of the protein:

Signal peptide:

amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 192-195, 225-228

N-myristoylation sites.

amino acids 42-47, 46-51, 136-141

FIGURE 139

CCTGGAGCCGGAAGCGCGGCTGCAGCAGGGCGAGGCTCCAGGTGGGGTCGGTTCCGCATCCAGCC
 TAGCGTGTCCACGATCGCGGCTGGGCTCCGGGACTTTTCGCTACCTGTTGCGTAGCGATCGAGGTGC
 TAGGGATCGCGGTCTTCCTTCGGGGATTCTTCCCGGCTCCCGTTCGTTCCCTCTGCCAGAGCGGAA
 CACGGAGCGGAGCCCCAGCGCCCGAACCCTCGGCTGGAGCCAGTTCCTAACTGGACCACGCTGCC
 ACCACCTCTCTTCAGTAAAGTTGTTATTGTTCTGATAGATGCCTTGAGAGATGATTTTGTGTTTG
 GGTCAAAGGGTGTGAAATTTATGCCCTACACAACCTACCTTGTGGAAAAAGGAGCATCTCACAGT
 TTTGTGGCTGAAGCAAAGCCACCTACAGTTACTATGCCTCGAATCAAGGCATTGATGACGGGGAG
 CCTTCCTGGCTTTGTCGACGTCACTAGGAACCTCAATTCTCTGCACTGCTGGAAGACAGTGTGA
 TAAGACAAGCAAAAGCAGCTGGAAAAAGAATAGTCTTTTATGGAGATGAAACCTGGGTAAATTA
 TTCCCAAAGCATTGTGTGAATATGATGGAACACCTCATTTTTCGTGTCAGATTACACAGAGGT
 GGATAATAATGTCACGAGGCATTGGATAAAGTATTAAGAGAGGAGATTGGGACATATTAATCC
 TCCACTACCTGGGGCTGGACCACATTGGCCACATTTTCAGGGCCCAACAGCCCCCTGATTGGGCAG
 AAGCTGAGCGAGATGGACAGCGTGCTGATGAAGATCCACACCTCACTGCAGTGAAGGAGAGAGA
 GACGCTTTTACCCAATTTGCTGGTTCTTTGTGGTGACCATGGCATGTCTGAAACAGGAAGTCACG
 GGGCTCTCTCCACCGAGGAGGTGAATACACCTCTGATTTTAATCAGTTCCTGCGTTTGAAGGAAA
 CCCGGTGATATCCGACATCCAAAGCACGTCCAAATAGACGGATGTGGCTGCGCACTGGCGATAGC
 ACTTGCGTTACCGATTCCAAAAGACAGTGTAGGGAGCCTCCTATTCCCAGTTGTGGAAGGAAGAC
 CAATGAGAGAGCAGTTGAGATTTTACATTTGAATACAGTGCAGCTTAGTAACTGTTGCAAGAG
 AATGTGCCGTCAATGAAAAAGATCCTGGGTTTGAGCAGTTTAAATGTCAGAAAGATTGCATGG
 GAACGTGATCAGACTGTACTTGGAGGAAAAGCATTGAGAAGTCTATTCAACCTGGGCTCCAAGG
 TTCTCAGGCAGTACCTGGATGCTCTGAAGACGCTGAGCTTGTCCCTGAGTGCACAAGTGGCCAG
 TTCTCACCCCTGCTCCTGCTCAGCGTCCCACAGGCACTGCACAGAAAGGCTGAGCTGGAAGTCCCA
 CTGTCACTCTCCTGGGTTTTCTCTGCTCTTTTATTGGTGATCCTGGTTCTTTTCGGCCGTTACGT
 CATTGTGTGCACCTCAGCTGAAAGTTCGTGCTACTTCTGTGGCTCTCGTGGCTGGCGGCAGGCT
 GCCTTTTCGTTTACCAGACTCTGGTTGAACACCTGGTGTGTGCCAAGTGTGGCAGTGGCCTGGAC
 AGGGGGCTCAGGGAAGGACGTGGAGCAGCCTTATCCCAGGCCTCTGGGTGTCCCAGACAGGTG
 TTCACATCTGTGCTGTGAGGTGAGTGCCTCAGTTCTTGAAAGCTAGGTTCTGCGACTGTTAC
 CAAGGTGATTGTAAAGAGCTGGCGGTACAGAGGAACAAGCCCCCAGCTGAGGGGGTGTGTGAA
 TCGGACAGCCTCCCAGCAGAGGTGTGGGAGCTGCAGCTGAGGGAAGAAGAGACAATCGGCCCTGGA
 CACTCAGGAGGGTCAAAAGGAGACTTGGTCGCACCACTCATCCTGCCACCCCCAGAATGCATCCT
 GCCTCATCAGGTCCAGATTTCTTTCAAGGCGGACGTTTTCTGTGGAATTCTTAGTCTTTGGCC
 TCGGACACCTTCATTGCTTAGCTGGGGAGTGGTGGTGAGGCAGTGAAGAAGAGGCGGATGGTCAC
 ACTCAGATCCACAGAGCCCAGGATCAAGGGACCCACTGCAGTGGCAGCAGGACTGTTGGGCCCCC
 ACCCAACCCCTGCACAGCCCTCATCCCTCTTGGCTTGAGCCGTGAGAGGCCCTGTGCTGAGTGT
 CTGACCGAGACACTCACAGCTTTGTCATCAGGGCACAGGCTTCCTCGGAGCCAGGATGATCTGTG
 CCACGCTTGCACCTCGGGCCCATCTGGGCTCATGCTCTCTCTCCTGCTATTGAATTAGTACCTAG
 CTGCACACAGTATGTAGTTACCAAAAGAATAAACGGCAATAATTGAGAAAAAAA

FIGURE 140

MRLGSGTFFATCCVAIEVLGIAVFLRGFFPAPVRSSARAEHGAEPPEPSAGASSNWTTLPPPLF
SKVVIVLIDALRDDVFVFGSKGVKEMPYTTYLVEKGASHSFVAEAKPPTVTMPRIKALMTGSLPGF
VDVIRNLNSPALLEDVIRQAKAAGKRIVFYGDETWVKLFPKHFEYDGTTSFFVSDYTEVDNNV
TRHLDKVLKRGDWDILILHYLGLDHHIGHISGPN SPLIGQKLSEMDSVLMKIHTSLQSKERETPLP
NLLVLCGDHGMSETGSHGASSTEEVNTPLILISSAFERKPGDIRHPKHVQ

Important features of the protein:**Signal peptide:**

amino acids 1-34

Transmembrane domain:

amino acids 58-76

N-glycosylation sites.

amino acids 56-60, 194-198

N-myristoylation sites.

amino acids 6-12, 52-58, 100-106, 125-131, 233-239, 270-276,
275-281, 278-284

Amidation site.

amino acids 154-158

Cell attachment sequence.

amino acids 205-208

FIGURE 141

GGCACGAGGCAAGCCTTCCAGGTTATCGTGACGCACCTTGAAAGTCTGAGAGCTACTGCCCTACA
GAAAGTTACTAGTGCCCTAAAGCTGGCGCTGGCACTGATGTTACTGCTGCTGTTGGAGTACAAC
TCCCTATAGAAAACAACCTGCCAGCACCTTAAGACCACTCACACCTTCAGAGTGAAGAACTTAAAC
CCGAAGAAATTCAGCATTTCATGACCAGGATCACAAAGTACTGGTCCTGGACTCTGGGAATCTCAT
AGCAGTTCAGATAAAAACTACATACGCCCAGAGATCTTCTTTGCATTAGCCTCATCCTTGAGCT
CAGCCTCTGCGGAGAAAGGAAGTCCGATTCTCCTGGGGGTCTCTAAAGGGGAGTTTTGTCTCTAC
TGTGACAAGGATAAAGGACAAAGTCATCCATCCCTTCAGCTGAAGAAGGAGAACTGATGAAGCT
GGCTGCCCCAAAAGGAATCAGCACGCCGGCCCTTCATCTTTTATAGGGCTCAGGTGGGCTCCTGGA
ACATGCTGGAGTCGGCGGCTCACCCGGATGGTTCATCTGCACCTCCTGCAATTGTAATGAGCCT
GTTGGGGTGACAGATAAATTTGAGAACAGGAAACACATTGAATTTTCATTTCACCCAGTTTGCAA
AGCTGAAATGAGCCCCAGTGAGGTCAGCGATTAGGAAACTGCCCCATTGAACGCCTTCCTCGCTA
ATTTGAACTAATTGTATAAAAACACCAAACCTGCTCACT

FIGURE 142

MLLLLLLEYNFPIENNCQHLKTTHTFRVKNLNPKKFSIHDQDHKVLVLDSGNLIAVPDKNYIRPEI
FFALASSLSSASAEKGSPILLGVSKGEFCLYCDKDKGQSHPSLQLKKEKLMKLAAQKESARRPFI
FYRAQVGSWNMLESAAHPGWFICTSCNCNEPVGVTDKFENRKHIEFSFQPVCKAEMSPSEVSD

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 33-36

N-myristoylation site.
amino acids 50-55, 87-92

Interleukin-1
amino acids 37-182

FIGURE 143

CTAGAGAGTATAGGGCAGAAGGATGGCAGATGAGTGACTCCACATCCAGAGCTGCCTCCCTTTAA
TCCAGGATCCTGTCTTCCTGTCTGTAGGAGTGCCTGTTGCCAGTGTGGGGTGAGACAAGTTTG
TCCCACAGGGCTGTCTGAGCAGATAAGATTAAGGGCTGGGTCTGTGCTCAATTAACCTGTGGG
CACGGGGGCTGGGAAGAGCAAAGTCAGCGGTGCCTACAGTCAGCACCATGCTGGGCCTGCCGTGG
AAGGGAGGTCTGTCTGGGCGCTGCTGCTGCTTCTCTTAGGCTCCCAGATCCTGCTGATCTATGC
CTGGCATTTCACGAGCAAAGGGACTGTGATGAACACAATGTCATGGCTCGTTACCTCCCTGCCA
CAGTGGAGTTTGCTGTCCACACATTCAACCAACAGAGCAAGGACTACTATGCCTACAGACTGGGG
CACATCTTGAATTCTGGAAGGAGCAGGTGGAGTCCAAGACTGTATTCTCAATGGAGCTACTGCT
GGGGAGAACTAGGTGTGGGAAATTTGAAGACGACATTGACAACTGCCATTTCCAAGAAAGCACAG
AGCTGAACAATACTTTCACCTGCTTCTTCACCATCAGCACCAGGCCCTGGATGACTCAGTTCAGC
CTCCTGAACAAGACCTGCTTGGAGGGATTCCACTTGAGTGAAACCCACTCACAGGCTTGTCCATGT
GCTGCTCCACATTCCGTGGACATCAGCACTACTCTCCTGAGGACTCTTCAGTGGCTGAGCAGCT
TTGGACTTGTTTGTATCTATTTTGCATGTGTTTGAGATCTCAGATCAGTGTTTTAGAAAATCC
ACACATCTTGAGCCTAATCATGTAGTGTAGATCATTAACATCAGCATTTTAAGAAAAAAAAAAAA
AA

$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}} \right) = \frac{\partial L}{\partial x}$

Important features of the protein:

amino acids 1-25

amino acids 117-121, 139-143

amino acids 9-15

FIGURE 145

CTGTGCAGCTCGAGGCTCCAGAGGCACACTCCAGAGAGAGCCAAGGTTCCTGACGCGATGAGGAAG
CACCTGAGCTGGTGGTGGCTGGCCACTGTCTGCATGCTGCTCTTCAGCCACCTCTCTGCGGTCCA
GACGAGGGGCATCAAGCACAGAATCAAGTGGAAACCGGAAGGCCCTGCCCAGCACTGCCCAGATCA
CTGAGGCCCAGGTGGCTGAGAACC GCCCGGAGCCTTCATCAAGCAAGGCCGCAAGCTCGACATT
GACTTCGGAGCCGAGGGCAACAGGTACTACGAGGCCAACTACTGGCAGTTCCCCGATGGCATCCA
CTACAACGGCTGCTCTGAGGCTAATGTGACCAAGGAGGCATTTGTCACCGGCTGCATCAATGCCA
CCCAGGCGGCGAACCAGGGGGAGTTCCAGAAGCCAGACAACAAGCTCCACCAGCAGGTGCTCTGG
CGGCTGGTCCAGGAGCTCTGCTCCCTCAAGCATTGCGAGTTTGGTTGGAGAGGGGCGCAGGACT
TCGGGTCAACATGCACCAGCCAGTGCTCCTCTGCCTTCTGGCTTTGATCTGGCTCATGGTGAAAT
AAGCTTGCCAGGAGGCTGGCAGTACAGAGCGCAGCAGCGAGCAAATCCTGGCAAGTGACCCAGCT
CTTCTCCCCCAAACCCACGCGTGTCTGAAGGTGCCCAGGAGCGGCGATGCACTCGCACTGCAAA
TGCCGCTCCCACGTATGCGCCCTGGTATGTGCCTGCGTTCTGATAGATGGGGGACTGTGGCTTCT
CCGTCACTCCATTCTCAGCCCCTAGCAGAGCGTCTGGCACACTAGATTAGTAGTAAATGCTTGAT
GAGAAGAACACATCAGGCACTGCGCCACCTGCTTCACAGTACTTCCCAACAACCTCTTAGAGGTAG
GTGTATTCCCGTTTTACAGATAAGGAACTGAGGCCCAGAGAGCTGAAGTACTGCACCCAGCATC
ACCAGCTAGAAAGTGGCAGAGCCAGGATTCAACCCCTGGCTTGTCTAACCCAGGTTTTCTGCTCT
GTCCAATTCCAGAGCTGTCTGGTGATCACTTTATGTCTCACAGGGACCCACATCCAAACATGTAT
CTCTAATGAAATTGTGAAAGCTCCATGTTTAGAAATAAATGAAAACACCTGA

FIGURE 146

MRKHLSSWWLATVCMLLFSLHSAVQTRGIKRIKWNRKALPSTAQITEAQVAENRPGAFIKQGRK
LDIDFGAEGNRYYEANYWQFPDGIHYNGCSEANVTKEAFVTGCINATQAAHQGEFQKPDNKLHQQ
VLWRLVQELCSLKHCEFWLERGAGLRVTMHQPVLLCLLALIWLMMVK

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 157-171

N-glycosylation sites.

amino acids 98-102, 110-114

Tyrosine kinase phosphorylation site.

amino acids 76-83

N-myristoylation sites.

amino acids 71-77, 88-94, 93-99, 107-113, 154-160

Amidation site.

amino acids 62-66

FIGURE 147

GCCTTGGCCTCCCAAAGGGCTGGGATTATAGGCGTGACCACCATGTCTGGTCCAGAGTCTCATTT
 CCTGATGATTTATAGACTCAAAGAAAACTATGTTTCAGAAGCTCTCTTCTCTTCTGGCCTCCTCT
 CTGTCTTCTTTCCCTCTTTCTTCTTATTTTAATTAGTAGCATCTACTCAGAGTCATGCAAGCTGG
 AAATCTTTCATTTTGCTTGTGAGTGGGGTAGGTCAGTCTTAGTTTTTATTTTTTGAAATTT
 CAACTTTCAGATTTCAGGGGGTACATGTGAAGGTTTGTTTTATGAGTATATTGCATTGATGTCTGAGG
 TTTGGGGT

FIGURE 148

MFRSSLLFWPPLCLLSLFLLLISSIYSECKLEIFHFACQWGRSLSLSFYFLKFQLSDSGGTCE
GLFYEYIA

Important features of the protein:

Signal peptide:

amino acids 1-25

N-myristoylation site.

amino acids 62-68

FIGURE 149

GTCTCCGCGTCACAGGAAGTTTACGACCCACAGGGCGGACAGCGCTCCCCTCTACCTGGAGACTTGAC
TCCCGCGCGCCCCAACCCCTGCTTATCCCTTGACCGTCGAGTGTGAGAGATCCTGCAGCCGCCAGTCC
CGGCCCCCTCTCCGCCCCACACCCACCCCTCCTGGCTCTTCCTGTTTTTACTCCTCCTTTTCATTGATA
ACAAAAGCTACAGCTCCAGGAGCCAGCGCGGGCTGTGACCCAAGCCGAGCGTGAAGAAATGGGGTT
CCTCGGGACCGGCACTTGGATTCTGGTGTTAGTGCTCCCGATTCAAGCTTTCCCCAACCTGGAGGAA
GCCAAGACAAATCTCTACATAATAGAGAATTAAGTGCAGAAAGACCTTTGAATGAACAGATTGCTGAA
GCAGAAGAAGACAAGATTAACAAAAACATATCTCCAGAAAACAAGCCAGGTCAGAGCAACTATTCTTT
TGTTGATAACTTGAACCTGCTAAAGGCAATAACAGAAAAGGAAAAAATTGAGAAAGAAAGACAATCTA
TAAGAAGCTCCCCACTTGATAATAAGTTGAATGTGGAAGATGTTGATTCAACCAAGAATCGAAAAGT
ATCGATGATTATGACTCTACTAAGAGTGGATTGGATCATAAATTTCAAGATGATCCAGATGGTCTTCA
TCACTAGACGGGACTCCTTTAACCCTGAAGACATTGTCCATAAAATCGCTGCCAGGATTATGAAG
AAAATGACAGAGCCGTGTTTGACAAGATTGTTTCTAACTACTTAATCTCGGCCTTATCAGAAAAGC
CAAGCACATACACTGGAAGATGAAGTAGCAGAGGTTTTACAAAAATTAATCTCAAAGGAAGCCAACAA
TTATGAGGAGGATCCCAATAAGCCACAAGCTGGACTGAGAATCAGGCTGGAAAAATACCAGAGAAA
TGACTCCAATGGCAGCAATTCAAGATGGTCTTGCTAAGGGAGAAAACGATGAAACAGTATCTAACACA
TTAACCTTGACAAATGGCTTGGAAGGAGAACTAAACCTACAGTGAAGACAACCTTTGAGGAACTCCA
ATATTTCCCAAATTTCTATGCGCTACTGAAAAGTATTGATTGAGAAAAGAAGCAAAAGAGAAAGAAA
CACTGATTACTATCATGAAAACACTGATTGACTTTGTGAAGATGATGCTGAAATATGGAACAATATCT
CCAGAAGAAGGTGTTTCCTACCTTGAAAACCTGGATGAAATGATTGCTCTTCAGACCAAAAAACAAGCT
AGAAAAAATGCTACTGACAATATAAGCAAGCTTTTCCAGCACCATCAGAGAAGAGTCATGAAGAAA
CAGACAGTACCAAGGAAGAAGCAGCTAAGATGGAAGGAATATGGAAGCTTGAAGGATTCCACAAAA
GATGATAACTCCAACCCAGGAGGAAAGACAGATGAACCCAAAGGAAAAACAGAAGCCTATTGGAAGC
CATCAGAAAAAATATTGAATGGTTGAAGAAACATGACAAAAAGGGAATAAAGAAGATTATGACCTTT
CAAAGATGAGAGACTTCATCAATAACAAGCTGATGCTTATGTGGAGAAAGGCATCCTTGACAAGGAA
GAAGCCGAGGCCATCAAGCGCATTTATAGCAGCCTGTAAAAAATGGCAAAAGATCCAGGAGTCTTTCAA
CTGTTTCAGAAAACATAATATAGCTTAAACACTTCTAATCTGTGATTAAAAATTTTTGACCCAAGG
GTTATTAGAAAGTGCTGAATTTACAGTAGTTAACCTTTTACAAGTGTTAAACATAGCTTCTTCCC
GTAAAACTATCTGAAAGTAAAGTTGTATGTAAGCTGAAAAAAGAAAAAAGAAAAA

FIGURE 150

MGFLGTGTWILVLVLP IQAFPKPGGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPPENKPG
QSNYSFVDNLNLLKAITEKEKIEKERQSIRSSPLDNKLNVEDVDSTKNRKLIDDYDSTKSGLDHK
FQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGLITESQAHTLEDEVAE
VLQKLISKEANNYEEDPNKPTSWTENQAGKIPKVTMAAIQDGLAKGENDETVSNTLTLTNGLE
RRTKTYSEDNFEELQYFPNFYALLKSIDSEKEAKEKETLITIMKTLIDFVKMMVKYGTISP EEGV
SYLENLDEMIALQTKNKLEKNATDNISKLFAPSEKSHEETDSTKEEAAKMEKEYGSLKDSTKDD
NSNPGGKTDEPKGKTEAYLEAIRKNIEWLKKHDKKGNKEDYDLSKMRDFINKQADAYVEKGILDK
EEAEA IKRIYSSL

N-glycosylation sites:

amino acids 68-71, 346-349, 350-353

Casein kinase II phosphorylation site:

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-
220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-
341, 369-372, 382-385, 386-389, 387-390

N-myristoylation sites:

amino acids 143-148, 239-244

FIGURE 151

CGGCTCGAGGCTCCCGCCAGGAGAAAGGAACATTCTGAGGGGAGTCTACACCTGTGGAGCTCAA
GATGGTCCTGAGTGGGGCGCTGTGCTTCCGAATGAAGGACTCGGCATTGAAGGTGCTTTATCTGC
 ATAATAACCAGCTTCTAGCTGGAGGGCTGCATGCAGGGAAGGTCAATTAAGGTGAAGAGATCAGC
 GTGGTCCCAATCGGTGGCTGGATGCCAGCCTGTCCCCGTCATCCTGGGTGTCCAGGGTGGAAG
 CCAGTGCCTGTATGTGGGGTGGGGCAGGAGCCGACTCTAACACTAGAGCCAGTGAACATCATGG
 AGCTCTATCTTGGTGCCAAGGAATCCAAGAGCTTCACCTTCTACCGGCGGGACATGGGGCTCACC
 TCCAGCTTCGAGTCGGCTGCCTACCCGGGCTGGTTCCTGTGCACGGTGCCTGAAGCCGATCAGCC
 TGTCACTCACCAGCTTCCCGAGAATGGTGGCTGGAATGCCCCATCACAGACTTCTACTTCC
 AGCAGTGTGACT**AGGGCA**ACGTGCCCCCAGAACTCCCTGGGCAGAGCCAGCTCGGGTGAGGGT
 GAGTGGAGGAGACCCATGGCGGACAATCACTCTCTCTGCTCTCAGGACCCACAGTCTGACTTAG
 TGGGCACCTGACCACTTTGTCTTCTGGTTCCAGTTTGGATAAATTCTGAGATTTGGAGCTCAGT
 CCACGGTCTCCCCACTGGATGGTGTCTACTGCTGTGGAACCTTGTA AAAACCATGTGGGGTAAA
 CTGGGAATAACATGAAAAGATTCTGTGGGGTGGGGTGGGGGAGTGGTGGGAATCATTCTGCT
 TAATGGTAAGTACAAGTGTACCCTGAGCCCCGAGGCCAACCCATCCCCAGTTGAGCCTTATA
 GGGTCAGTAGCTCTCCACATGAAGTCTGTCACTCACCAGTGTGCAGGAGAGGGAGGTGGTCATA
 GAGTCAGGGATCTATGGCCCTTGCGCCAGCCCCACCCCTTCCCTTTAATCCTGCCACTGTCATA
 TGCTACCTTTCCATCTCTTCCCTCATCATCTTGTGTGGGCATGAGGAGGTGGTGTATGTCAGAA
 GAAATGGCTCGAGCTCAGAAGATAAAAGATAAGTAGGGTATGCTGATCCTTTTTAAAAACCAA
 GATACAATCAAAATCCCAGATGCTGGTCTCTATTCCCATGAAAAAGTGCTCATGACATATTGAGA
 AGACCTACTTACAAAGTGGCATATATTGCAATTTATTTTAAATTAAGATACCTATTTATATATT
 TCTTTATAGAAAAAGTCTGGAAGAGTTTACTTCAATTGTAGCAATGTCAAGGTGGTGGCAGTAT
 AGGTGATTTTTCTTTTAATTCTGTAAATTTATCTGTATTTCCTAATTTTCTACAATGAAGATGA
 ATTCTTGTATAAAAAATAAGAAAAAGAAATTAATCTTGAGGTAAGCAGAGCAGACATCATCTCTGA
 TTGTCTCAGCCTCCACTTCCCCAGAGTAAATTCAAATTTGAATCGAGCTCTGCTGCTCTGGTTGG
 TTGTAGTAGTGATCAGGAAACAGATCTCAGCAAAGCCACTGAGGAGGAGGCTGTGCTGAGTTTGT
 GTGGCTGGAATCTCTGGGTAAGGAACCTTAAAGAACAAAAATCATCTGGTAATCTTTCCTAGAAG
 GATCACAGCCCCCTGGGATTCCAAGGCATTGGATCCAGTCTCTAAGAAGGCTGCTGTACTGGTTGA
 ATTTGTGTCCTTCAAATTCACATCCTTCTTGGAAATCTCAGTCTGTGAGTTTATTTGGAGATAAG
 GTCTCTGCAGATGTAGTTAGTTAAGACAAGGTCATGCTGGATGAAGGTAGACCTAAATTCATAT
 GACTGGTTTCCTTGATGAAAAGGAGAGGACACAGAGACAGAGGAGACGCGGGGAAGACTATGTA
 AAGATGAAGGCAGAGATCGGAGTTTTCAGCCACAAGCTAAGAAACACCAAGGATTGTGGCAACC
 ATCAGAAGCTTGGAAGAGGCAAGAAGAATTCTTCCCTAGAGGCTTTAGAGGGATAACGGCTCTG
 CTGAAACCTTAATCTCAGACTTCCAGCCTCCTGAACGAAGAAAGATAAATTCGGCTGTTTTAA
 GCCACCAAGGATAATTGGTTACAGCAGCTCTAGGAACTAATACAGCTGCTAAAATGATCCCTGT
 CTCCTCGTGTTTACATTCTGTGTGTGTCCCCCTCCCAATGTACCAAAGTTGTCTTTGTGACCAA
 TAGAATATGGCAGAAGTGATGGCATGCCACTTCCAAGATTAGGTTATAAAAGACACTGCAGCTTC
 TACTTGAGCCCTCTCTCTGCCCCACCGCCCCCAATCTATCTTGGCTCACTCGCTCTGGGGG
 AAGCTAGCTGCCATGCTATGAGCAGGCCTATAAAGAGACTTACGTGGTAAAAAATGAAGTCTCCT
 GCCCACAGCCACATTAGTGAACCTAGAAGCAGAGACTCTGTGAGATAATCGATGTTTGTGTTTT
 AAGTTGCTCAGTTTGGTCTAACTTGTTATGCAGCAATAGATAAATAATATGCAGAGAAAGAG

FIGURE 152

MVLSGALCFRMKDSALKVLYLHNNQLLAGGLHAGKVIKGEESISVVPNRWLDASLSPVILGVQGGG
 QCLSCGVGQEPTLTLEPVNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGWFLCTVPEADQP
 VRLTQLPENGGWNAIPITDFYFQQCD

N-myristoylation sites.

amino acids 29-34, 30-35, 60-65, 63-68, 73-78, 91-96, 106-111

Interleukin-1 signature.

amino acids 111-131

Interleukin-1 proteins.

amino acids 8-29, 83-120, 95-134, 64-103

FIGURE 153

CTTCAGAACAGGTTCTCCTTCCCCAGTCACCCAGTTGCTCGAGTTAGAATTGTCTGCAATGGCCGC
 CCTGCAGAAATCTGTGAGCTCTTCCCTTATGGGGACCCTGGCCACCAGCTGCCTCCTTCTTTGG
 CCCTCTTGGTACAGGGAGGAGCAGCTGCGCCCATCAGCTCCCACTGCAGGCTTGACAAGTCCAAC
 TTCCAGCAGCCCTATATCACCAACCGCACCTTCATGCTGGCTAAGGAGGCTAGCTTGGCTGATAA
 CAACACAGACGTTTCGTCTCATTGGGGAGAACTGTTCCACGGAGTCAGTATGAGTGAGCGCTGCT
 ATCTGATGAAGCAGGTGCTGAACTTCACCCTTGAAGAAGTGCTGTTCCCTCAATCTGATAGGTTT
 CAGCCTTATATGCAGGAGGTGGTGCCTTCCCTGGCCAGGCTCAGCAACAGGCTAAGCACATGTCA
 TATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGCAAAAGCTGAAGGACACAGTGAAAAAGC
 TTGGAGAGAGTGGAGAGATCAAAGCAATTGGAGAACTGGATTGCTGTTTATGTCTCTGAGAAAT
 GCCTGCATTTGACCAGAGCAAAGCTGAAAAATGAATAACTAACCCCTTCCCTGCTAGAAATAA
 CAATTAGATGCCCCAAAGCGATTTTTTTTAAACCAAAGGAAGATGGGAAGCCAACTCCATCATG
 ATGGGTGGATTCCAAATGAACCCCTGCGTTAGTTACAAAGGAAACCAATGCCACTTTTGTATATA
 AGACCAGAAGGTAGACTTTCTAAGCATAGATATTTATTGATAACATTTTCATTGTAAGTGGTGTTT
 TATACACAGAAAAACAATTTATTTTAAATAATTGTCTTTTCCATAAAAAAGATTACTTTCCAT
 TCCTTTAGGGGAAAAAACCCCTAAATAGCTTCATGTTCCATAATCAGTACTTTATATTTATAAA
 TGTATTTATTATTATTATAAGACTGCATTTTATTATATCATTTTATTAATATGGATTATTTAT
 AGAAACATCATTTCGATATTGCTACTTGAGTGTAAGGCTAATATTGATATTTATGACAATAATTAT
 AGAGCTATAACATGTTTATTTGACCTCAATAAACACTTGGATATCCC

FIGURE 154

MAALQKSVSSFLMGTLATSCLLLLALLVQGGAAAPISSHCRLDKSNFQQPYITNRTFMLAKEASL
ADNNTDVRLIGEKLFGVSMSERCYLMKQVLNFTLEEVLFQSDRFQPYMQEVVPFLARLSNRLS
TCHIEGDDLHIQRNVQKLKDTVKKLGESGEIKAIGELDLEFMSLRNACI

Important features of the protein:

Signal peptide:

amino acids 1-33

N-glycosylation sites.

amino acids 54-58, 68-72, 97-101

N-myristoylation sites.

amino acids 14-20, 82-88

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 10-21

FIGURE 155

GGCTTGCTGAAAATAAAATCAGGACTCCTAACCTGCTCCAGTCAGCCTGCTTCCACGAGGCCTGT
CAGTCAGTGCCCGACTTGTGACTGAGTGTGCAGTGCCAGCATGTACCAGGTCAGTGCAGAGGGC
TGCCTGAGGGCTGTGCTGAGAGGGAGAGGAGCAGAGATGCTGCTGAGGGTGGAGGGAGGCCAAGC
TGCCAGGTTTGGGGCTGGGGGCCAAGTGGAGTGAGAACTGGGATCCCAGGGGGAGGGTGCCAGAT
GAGGGAGCGACCCAGATTAGGTGAGGACAGTTCTCTCATTAGCCTTTTCCCTACAGGTGGTTGCAT
TCTTGGCAATGGTCATGGGAACCCACACCTACAGCCACTGGCCAGCTGCTGCCCCAGCAAAGGG
CAGGACACCTCTGAGGAGCTGCTGAGGTGGAGCACTGTGCCTGTGCCTCCCCTAGAGCCTGCTAG
GCCCCAACCGCCACCCAGAGTCCTGTAGGGCCAGTGAAGATGGACCCCTCAACAGCAGGGCCATCT
CCCCCTGGAGATATGAGTTGGACAGAGACTTGAACCGGCTCCCCAGGACCTGTACCAGCCCCGT
TGCCTGTGCCCCGACTGCGTCAGCCTACAGACAGGCTCCCACATGGACCCCCGGGGCAACTCGGA
GCTGCTCTACCACAACCAGACTGTCTTCTACAGGCGGCCATGCCATGGCGAGAAGGGCACCCACA
AGGGCTACTGCCTGGAGCGCAGGCTGTACCGTGTTCCTTAGCTTGTGTGTGTGTGCGGGCCCGT
GTGATGGGGCTAGCCGGACCTGCTGGAGGCTGGTCCCTTTTGGGAAACCTGGAGCCAGGTGTACA
ACCCTTGCCATGAAGGGCCAGGATGCCCAGATGCTTGGCCCCCTGTGAAGTGCTGTCTGGAGCAG
CAGGATCCCGGGACAGGATGGGGGGCTTTGGGGAAAACCTGCACTTCTGCACATTTTGAAAAGAG
CAGCTGCTGCTTAGGGCCCGCGAAGCTGGTGTCTGTCAATTTCTCTCAGGAAAGGTTTTCAAA
GTTCTGCCCCATTTCTGGAGGCCACCACTCCTGTCTCTTCTCTTTTCCCATCCCCTGCTACCCCTG
GCCCAGCACAGGCACTTTCTAGATATTTCCCCCTTGCTGGAGAAGAAAGAGCCCTGGTTTTATT
TGTTTGTGTTTACTCATCACTCAGTGAGCATCTACTTTGGGTGCATTCTAGTGTAGTTACTAGTCTT
TTGACATGGATGATTCTGAGGAGGAAGCTGTTATTGAATGTATAGAGATTTATCCAAATAAATAT
CTTTATTTAAAAATGAAAAA

FIGURE 156

MRERPRLGEDSSLISLFLQVVAFLAMVMGHTYSHWPSCCPKQDTSEELLRWSTVPVPPLEPA
RPNRHPESCRASEDGPLNSRAISPWRYELDRDLNRLPQDLYHARCLCPHCVSLQTGSHMDPRGNS
ELLYHNQTVFYRRPCHGEKGTHKGYCLERRLYRVSLACVCVRPRVMG

Important features of the protein:

Signal peptide:

amino acids 1-32

N-glycosylation site.

amino acids 136-140

Tyrosine kinase phosphorylation site.

amino acids 127-135

N-myristoylation sites.

amino acids 44-50, 150-156

FIGURE 157

CCGGCGATGTCGCTCGTGTGCTAAGCCTGGCCGCGCTGTGCAGGAGCGCCGTACCCCGAGAGCC
GACCGTTCAATGTGGCTCTGAACTGGGCCATCTCCAGAGTGGATGCTACAACATGATCTAATCC
CCGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACAACACTAGTGTGCAACAGGGGACTATTCA
ATTTTGATGAATGTAAGCTGGGTACTCCGGGCAGATGCCAGCATCCGCTTGTTGAAGGCCACCAA
GATTTGTGTGACGGGCAAAGCAACTTCCAGTCCTACAGCTGTGTGAGGTGCAATTACACAGAGG
CCTTCCAGACTCAGACCAGACCCTCTGGTGGTAAATGGACATTTTCCTACATCGGCTTCCTGTGA
GAGCTGAACACAGTCTATTTTCATTGGGGCCATAATATTCCTAATGCAAATATGAATGAAGATGG
CCCTTCCATGTCTGTGAATTCACCTCACCAGGCTGCCTAGACCACATAATGAAATATAAAAAAA
AGTGTGTCAAGGCCGGAAGCCTGTGGGATCCGAACATCACTGCTTGTGAAGAAGATGAGGAGACA
GTAGAAGTGAAGTTCACAACCACTCCCTGGGAAACAGATACATGGCTCTTATCCAACACAGCAC
TATCATCGGGTTTCTCAGGTGTTTGAAGGACACACAGAAAGAAACAAACGCGAGCTTCAGTGGTGA
TTCCAGTGACTGGGGATAGTGAAGGTGCTACGGTGCAGCTGACTCCATATTTTCCTACTTGTGGC
AGCGACTGCATCCGACATAAAGGAACAGTTGTGCTCTGCCCACAAACAGGCGTCCCTTTCCCTCT
GGATAACAACAAAAGCAAGCCGGGAGGCTGGCTGCCTCTCCTCTGCTGTCTCTGCTGGTGGCCA
CATGGGTGCTGGTGGCAGGGATCTATCTAATGTGGAGGCACGAAAGGATCAAGAAGACTTCCTTT
TCTACCACCACACTACTGCCCCCCATTAAGGTTCTTGTGGTTTACCCATCTGAAATATGTTTCCA
TCACACAATTTGTTACTTCACTGAATTTCTTCAAACCATTCGAGAAGTGAGGTGATCCTTGAAA
AGTGGCAGAAAAAGAAAATAGCAGAGATGGGTCCAGTGCAGTGGCTTGCCACTCAAAGAAGGCA
GCAGACAAAGTCGTCTTCCTTCTTTCCAATGACGTCAACAGTGTGTGCGATGGTACCTGTGGCAA
GAGCGAGGGCAGTCCCAGTGAGAACTCTCAAGACCTCTTCCCCCTTGCTTTAACCTTTTCTGCA
GTGATCTAAGAAGCCAGATTCTCTGCACAAATACGTGGTGGTCTACTTTAGAGAGATTGATACA
AAAGACGATTACAATGCTCTCAGTGTCTGCCCCAAGTACCACCTCATGAAGGATGCCACTGCTTT
CTGTGCAGAACTTCTCCATGTCAAGCAGCAGGTGTCAGCAGGAAAAAGATCACAAGCCTGCCACG
ATGGCTGCTGCTCCTTGTAG

8

FIGURE 158

MSLVLLSLAALCRSAVPREPTVQCGSETGPSPEWMLQHDLPDGLRDLRVEPVTTSVATGDYSILMNVS
 LRADASIRLLKATKICVTGKSNFQSYSCVRCNYTEAFQTQTRPSGGKWTFSYIGFPVELNTVYFIGAHNIP
 NANMNEDGPSMSVNFTSPGCLDHIMKYKKKCVKAGSLWDPNITACKKNEETVEVNFTTTPLGNRYMALIQH
 STIIGFSQVFEPHQKKQTRASVVIPVTGDSEGATVQLTPYFPTCGSDCIRHKGTVVLCPTGVPFPLDNNK
 SKPGGWLPLLLLSLLVATWVLVAGIYLMWRHERIKKTSFSTTTLLPPIKVLVVYPSEICFHHTICYFTEFL
 QNHCRSEVILEKWQKKKIAEMGPVQWLATQKKAADKVVFLSNDVNSVCDGTCGKSEGSPSENSQDLFPLA
 FNLFCSDLRSQIHLHKYVVVYFREIDTKDDYNALSVC PKYHLMKDATAFCAELLHV KQVVSAGKRSQACHD
 GCCSL

Important features of the protein:

Signal peptide:

amino acids 1-14

Transmembrane domain:

amino acids 290-309

N-glycosylation sites.

amino acids 67 - 71, 103 - 107, 156 - 160, 183 - 187, 197 - 201 and 283
 - 287

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 228 - 232 and 319 - 323

Casein kinase II phosphorylation sites.

amino acids 178 - 182, 402 - 406, 414 - 418 and 453 - 457

N-myristoylation site.

amino acids 116-122

Amidation site.

amino acids 488-452

FIGURE 159

AGCCACCAGCGCAACATGACAGTGAAGACCCTGCATGGCCCAGCCATGGTCAAGTACTTGCTGCT
GTCGATATTGGGGCTTGCCTTTCTGAGTGAGCGGCAGCTCGGAAAATCCCCAAAGTAGGACATA
CTTTTTTCCAAAAGCCTGAGAGTTGCCCCCTGTGCCAGGAGGTAGTATGAAGCTTGACATTGGC
ATCATCAATGAAAACCAGCGCGTTTCCATGTCACGTAACATCGAGAGCCGCTCCACCTCCCCCTG
GAATTACACTGTCACTTGGGACCCCAACCGGTACCCCTCGGAAGTTGTACAGGCCCAGTGTAGGA
ACTTGGGCTGCATCAATGCTCAAGGAAAGGAAGACATCTCCATGAATCCGTTCCCATCCAGCAA
GAGACCCTGGTCGTCCGGAGGAAGCACCAAGGCTGCTCTGTTTCTTTCCAGTTGGAGAAGGTGCT
GGTGACTGTTGGCTGCACCTGCGTCACCCCTGTCATCCACCATGTGCAGTTAAGAGGTGCATATCC
ACTCAGCTGAAGAAG

FIGURE 16o

MTVKTLHGPMVKYLLLSILGLAFLSEAAARKIPKVGHTFFQKPESCPPVPGGSMKLDIGIINEN
 QRVSMERNIESRSTSPWNYTWTWDPNRYPPSEVVQAQCRNLGCINAQKEDISMNSVPIQQETLVV
 RRKHQGCSSVSFQLEKVLVTVGCTCVTPVIHHVQ

Signal sequence:

amino acids 1-30

N-glycosylation site.

amino acids 83-87

N-myristoylation sites.

amino acids 106-111, 136-141

FIGURE 161

AACTGGCCAAACAAAAACGAAAGCACTCCGTGCTGGAAGTAGGAGGAGAGTCAGGACTCCCAGG
 ACAGAGAGTGCACAACTACCCAGCACAGCCCCCTCCGCCCCCTCTGGAGGCTGAAGAGGGATT
 CAGCCCCGTGCCACCCACAGACACGGGGCTGACTGGGGTGTCTGCCCCCTTGGGGGGGGGACAGC
 AGGGCCTCAGGCCTGGGTGCCACCTGGCACCTAGAAGATGCTGTGCCCTGGTTCTTGCTGTCCT
 TGGCACTGGGGCCGAAGCCAGTGGTCCCTTCTCTGGAGAGGCTTGTGGGGCTCAGGACGCTACC
 CACTGCTCTCCGGGCTCTCCTGCCGCTCTGGGACAGTGACATACTCTGCCTGCCTGGGGACAT
 CGTGCCTGCTCCGGGCCCCGTGCTGGCGCTACGCACCTGCAGACAGAGCTGGTGTGAGGTGCC
 AGAAGGAGACCGACTGTGACCTCTGTCTGCGTGTGGCTGTCCACTTGGCCGTGCATGGGCACTGG
 GAAGAGCCTGAAGATGAGGAAAAGTTTGGAGGAGCAGCTGACTCAGGGGTGGAGGAGCCTAGGAA
 TGCTCTCTCCAGGCCCCAAGTCGTGCTCTCCTTCCAGGCCTACCCTACTGCCCGCTGCGTCTCTGC
 TGGAGGTGCAAGTGCTGCTGCCCCCTTGTGCACTTTGGTCACTCTGTGGGCTCTGTGGTATATGAC
 TGCTTCGAGGCTGCCCTAGGGAGTGAGGTACGAATCTGGTCTTATACTCAGCCCCAGGTACGAGAA
 GGAACCTAACACACACAGCAGCTGCCCTGCCCTGGCTCAACGTGTACAGCAGATGGTGACA
 ACGTGCTCTGGTTCTGAATGTCTCTGAGGAGCAGCACTTCGGCCTCTCCCTGTACTGGAATCAG
 GTCCAGGGCCCCCAAACCCCGGTGGCACAAAACCTGACTGGACCGCAGATCATTACCTTGAA
 CCACACAGACCTGGTTCCCTGCCTCTGTATTAGGTGTGGCCTCTGGAACCTGACTCCGTTAGGA
 CGAATCTGCCCCCTCAGGGAGGACCCCCGCGCACACCAGAACCTCTGGCAAGCCGCCCCGACTG
 CGACTGCTGACCCTGCAGAGCTGGCTGCTGGACGCACCGTGCTCGCTGCCCGCAGAAGCGGCACT
 GTGCTGGCGGGCTCCGGGTGGGGACCCCTGCCAGCCACTGGTCCCACCGCTTCTCTGGGAGAACG
 TCACTGTGGACAAGTTCTCGAGTTCCTTCTGCTGAAAGGCCACCTAACCTCTGTGTTAGGTG
 AACAGCTCGGAGAAGCTGCAGCTGCAGGAGTGCTTGTGGGCTGACTCCCTGGGGCCTCTCAAAGA
 CGATGTGCTACTGTTGGAGACACGAGGCCCCCAGGACAACAGATCCCTCTGTGCTTGGAAACCCA
 GTGGCTGTACTTCACTACCCAGCAAAGCCTCCACGAGGGCAGCTCGCCTTGGAGAGTACTTACTA
 CAAGACCTGCAGTCAGGCCAGTGTCTGCAGCTATGGGACGATGACTTGGGAGCGCTATGGGCCTG
 CCCCATGGACAAATACATCCACAAGCGCTGGGCCCTCGTGTGGCTGGCCTGCCTACTCTTTGCCG
 CTGCGCTTTCCCTCATCCTCCTTCTCAAAAAGGATCACGCGAAAGGGTGGCTGAGGCTCTTGAAA
 CAGGACGTCCGCTCGGGGGCGGCCGCCAGGGGCCGCGGGCTCTGCTCCTACTCAGCCGATGA
 CTCGGGTTTCGAGCGCTGGTGGGCGCCCTGGCGTGGCCCTGTGCCAGCTGCCGCTGCGCGTGG
 CCGTAGACCTGTGGAGCCGTGCTGAACTGAGCGCGCAGGGGGCCCGTGGCTTGGTTTACGCGCAG
 CGGCGCCAGACCTGCAGGAGGGCGGCGTGGTGGTCTTGTCTTCTCTCCCGGTGCGGTGGCGCT
 GTGCAGCGAGTGGCTACAGGATGGGGTGTCCGGGCCCGGGGCGCACGGCCCGCACGACGCCTTCC
 GCGCCTCGCTCAGCTGCGTGCTGCCCCGACTTCTTGACGGGCCGGGCGCCCGGAGCTACGTGGGG
 GCCTGCTTCGACAGGCTGCTCCACCCGACGCCGTACCCGCCCTTTCCGCACCGTGCCCGTCTT
 CACACTGCCCTCCCAACTGCCAGACTTCTTGGGGGCCCTGCAGCAGCCTCGCGCCCCGCGTTCGG
 GGCGGCTCCAAGAGAGAGCGGAGCAAGTGTCCTGGGGCCCTTACGCCAGCCCTGGATAGCTACTTC
 CATCCCCGGGGACTCCCGCGCGGGGACGCGGGGTGGGACCAGGGCGGGACCTGGGGCGGGGGA
 CGGGACTTAAATAAAGGCAGACGCTGTTTTCTAAAAAA

FIGURE 162

MPVPWFLLSLALGRSPVVLSELERLVGPQDATHCSPGLSCLWDS DILCLPGDIVPAPGPVLAPTHLQTELV
 LRCQKETDCDLCLRVAVHLAVHGHWEPEDEEKFGGAADSGVEEPRNASLQAQVVL SFQAYPTARCVLLEV
 QVPAALVQFGQSVG SVVYDCFEAALGSEVRIWSYTQPRYEKEINHTQQLPALPWLNV SADGDNVHLVNLVS
 EEQHFGLSLYWNQVQGPPKPRWHKNLTGPQIITLNHTDLVPCLCIQVWLEPDSVRTNICPFREDPRAHQN
 LWQAARLRLTLQSWLLDAPCSLPAAEALCWRAFGGDFCQPLVPPLSWENVTVDKVLEFELLKGHPNLCVQ
 VNSSEKLQEQECLWADSLGPKDDVLLLETRGPQDNRS LCALEPSGCTSLPSKASTRAARLGEYLLQDLQS
 GQCLQLWDDDLGALWACPMCKYIHKRWALVWLACLLFAAALSLILLKKDHAKGWLRLKQDVRSGAAARG
 RAALLYSADDSGFERLVGALASALCQLPLRVAVDLWSRRELSAQGPVAVFHAQRRQTLQEGGVVLLFSP
 GAVALCSEWLQDGVSGPGAHGPHDAFRASLSCVLPDFLQGRAPGSYVGACFDRLHPDAVPALFRTVPVET
 LPSQLPDFLGALQQPRAPRSGRLQERAEQVSRALQPALDSYFHPPGTPAPGRGVGPGAGPGAGDGT

Signal sequence:

amino acids 1-20

Transmembrane domain.

amino acids 453-475

N-glycosylation sites.

amino acids 118-121, 186-189, 198-201, 211-214, 238-241, 248-251,
 334-337, 357-360, 391-394

Glycosaminoglycan attachment site.

amino acids 583-586

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 552-555

N-myristoylation sites.

amino acids 107-112, 152-157, 319-324, 438-443, 516-521, 612-617,
 692-697, 696-701, 700-705

FIGURE 163

GGGAGGGCTCTGTGCCAGCCCCGATGAGGACGCTGCTGACCATCTTGACTGTGGGATCCCTGGCT
 GCTCACGCCCCCTGAGGACCCCTCGGATCTGCTCCAGCACGTGAAATTCAGTCCAGCAACTTTGA
 AAACATCCTGACGTGGGACAGCGGGCCAGAGGGCACCCAGACACGGTCTACAGCATCGAGTATA
 AGACGTACGGAGAGAGGGGACTGGGTGGCAAAGAAGGGCTGTCAGCGGATCACCCGGAAGTCTGC
 AACCTGACGGTGGAGACGGGCAACCTCACGGAGCTCTACTATGCCAGGGTCACCGCT
 GTCAGTGCGGGAGGCCGGTCAGCCACCAAGATGACTGACAGGTTGAGTCTCTGTCAGCACACTAC
 CCTCAAGCCACCTGATGTGACCTGTATCTCCAAAGTGAGATCGATTGAGATGATTGTTTCATCCTA
 CCCCCAGCCAATCCGTGCAGGCGATGGCCACCGGCTAACCTGGAAGACATCTCCATGACCTG
 TTCTACCACTTAGAGCTCCAGGTCAACCGCACCTACCAAATGCACCTTGGAGGGAAGCAGAGAGA
 ATATGAGTTCTTCGGCCTGACCCCTGACACAGAGTTCTTGGCACCATCATGATTGCGTTCCCA
 CCTGGGCCAAGGAGAGTGCCCCCTACATGTGCCGAGTGAAGACACTGCCAGACCGGACATGGACC
 TACTCCTTCTCCGGAGCCTTCTGTCTCCATGGGCTTCTCGTCGAGTACTCTGCTACCTGAG
 CTACAGATATGTACCAAGCCGCTGCACCTCCCAACTCCCTGACGTCAGCGAGTCCCTGAGCTT
 TCCAGCCGCTGCGCTTCATCCAGGAGCAGTCTGATCCCTGTCTTTGACCTCAGCGGCCCCAGC
 AGTCTGGCCAGCCTGTCCAGTACTCCAGATCAGGGTGTCTGGACCCAGGGAGCCCGCAGGAGC
 TCCACAGCGGCATAGCCTGTCCGAGATCACCTACTTAGGGCAGCCAGACATCTCCATCCTCCAGC
 CCTCCACGTGCCACCTCCCCAGATCCTCTCCCCACTGTCTATGCCCCAACGCTGCCCTGAG
 GTCGGGCCCCCATCCTATGCACCTCAGGTGACCCCGAAGCTCAATTCCATTCTACGCCCCACA
 GGCCATCTCTAAGGTCCAGCCTTCTCCTATGCCCCCTCAAGCCACTCCGGACAGCTGGCCTCCCT
 CCTATGGGGTATGCAAGGTTCTGGCAAAGACTCCCCCACTGGGACACTTCTAGTCTCTAAA
 CACCTTAGGCCTAAAGGTCAGCTTCAGAAAGAGCCACCAGCTGGAAGCTGCATGTTAGGTGGCCT
 TTCTCTGCAGGAGGTGACCTCCTTGCTATGGAGGAATCCCAAGAAGCAAAATCATTGCACAGC
 CCTGGGGATTGTCACAGACAGAACATCTGACCCAAATGTGCTACACAGTGGGGAGGAAGGGACA
 CCACAGTACCTAAAGGGCCAGCTCCCCCTCCTCTCCTCAGTCCAGATCGAGGGCCACCCCATGTC
 CCTCCCTTTGCAACCTCCTTCCGGTCCATGTTCCCCCTCGGACCAAGGTCCAAGTCCCTGGGGCC
 TGCTGGAGTCCCTTGTGTGTCCCAAGGATGAAGCCAAGAGCCAGCCCTGAGACCTCAGACCTG
 GAGCAGCCACAGAACTGGATTCTCTTTTTCAGAGGCTGGCCCTGACTGTGCAGTGGGAGTCTTG
 AGGGGAATGGGAAAGGCTTGGTGCTTCTCCCTGTCCCTACCCAGTGTACATCCTTGGCTGTCA
 ATCCCATGCTGCCCCATGCCACACTCTGCGATCTGGCCTCAGACGGGTGCCCTTGAGAGAAGC
 AGAGGGAGTGGCATGCAGGGCCCCTGCCATGGGTGCGCTCCTCACCGGAACAAAGCAGCATGATA
 AGGACTGCAGCGGGGGAGCTCTGGGGAGCAGCTTGTGTAGACAAGCGCGTGTGCTGAGCCCTG
 CAAGGCAGAAATGACAGTGCAAGGAGGAAATGCAGGGAACTCCCGAGGTCCAGAGCCCCACCTC
 CTAACACCATGGATTCAAAGTGCTCAGGGAATTTGCCCTCTCCTTGCCCCATTCTTGCCAGTTTC
 ACAATCTAGCTCGACAGAGCATGAGGCCCCCTGCCCTCTTCTGTCAATTGTTCAAAGGTGGGAAGAGA
 GCCTGGAAAAGAACCAGGCCTGGAAAAGAACCAGAAGGAGGCTGGGCAGAACCAGAACAACCTGC
 ACTTCTGCCAAGGCCAGGGCCAGCAGGACGGCAGGACTCTAGGGAGGGGTGTGGCCTGCAGTCA
 TTCCCAGCCAGGGCACTGCCTGACGTTGCACGATTTGAGCTTCATTCTCTGATAGAACAAAGC
 GAAATGCAGGTCCACCAGGGAGGGAGACACACAAGCCTTTTCTGCAGGCAGGAGTTTCAGACCCCT
 ATCCTGAGAATGGGGTTTGAAGGAAGGTGAGGGCTGTGGCCCTGGACGGGTACAATAACACAC
 TGTACTGATGTACAACTTTGCAAGCTCTGCCCTGGGTTTACGCCATCTGGGCTCAAATTCAGC
 CTCACCACTCACAAGCTGTGTGACTTCAAACAAATGAAATCAGTCCCCAGAACCTCGGTTTCTCTC
 ATCTGTAATGTGGGGATCATAACCTACCTCATGGAGTTGTGGTGAAGATGAAATGAAGTCATG
 TCTTTAAAGTGCTTAATAGTGCCTGGTACATGGGCAGTGGCCAATAACGGTAGCTATTTAAAAA
 AAAAAAA

FIGURE 164

MRTLLTILTVGSLAAHAPEDPSDLLQHVKFQSSNFENILTWDSGPEGTPDVTYSIEYKTYGERDW
 VAKKGCQRITRKSCNLTVETGNLTELYARVTAVSAGGRSATKMTDRFSSLQHTTLKPPDVTCTIS
 KVRSIQMIVHPTPTPIRAGDGHRLTLEDIFHDLFYHLELQVNRTYQMHLGGKQREYEFFGLTPDT
 EFLGTIMICVPTWAKESAPYMCRVKTLPDRTWTYSFSGAFLFSMGFLVAVLCYLSYRYVTKPPAP
 PNSLVQRVLTFQPLRFIQEHVLIPVFDLSGPSSLAQPVQYSQIRVSGPREPAGAPQRHSLSEIT
 YLGQPDISILQPSNVPPPQILSPLSYAPNAAPEVGPSPSYAPQVTPEAQFPFYAPQAISKVQPSSY
 APQATPDSWPPSYGVCMEGSGKDSPTGTLSSPKHLRPGQLQKEPPAGSCMLGGLSLQEVTSIAM
 EESQEAKSLHQPLGICTDRTSDPNVLHSGEETPQYLKGQLPLLSSVQIEGHPMSLPLQPPSGPC
 SPSPDQGPSPWGLLESIVCPKDEAKSPAPETSDLEQPTELDSLFRGLALTVQWES

Signal sequence.

amino acids 1-17

Transmembrane domain.

amino acids 233-250

N-glycosylation sites.

amino acids 80-83, 87-90, 172-175

N-myristoylation sites.

amino acids 11-16, 47-52, 102-107, 531-536, 565-570

FIGURE 166

MAAAPGLLFWLFLVGLALWWVPGQSDLSHGRRFSDLKVCGDEECSMLMYRGKALEFTGPDCRFVN
 FKKGDDVYVYYKLAGGSLELWAGSVEHSFGYFPKDLIKVLHKYTEELHIPAETDFVCFEGGRD
 DFNSYNVEELLGSLELEDSVPEESKKAEEVSQHREKSPEESRGRELDPVPEPEAFRAEDSEGEA
 FSESTGLQGQPSAQESHPTSGPAANAQGVQSSLDTFEEILHDKLKVPGSESRTGNSSPASVER
 EKTDAYKVLKTEMSQRGSGQCVIHYSKGFRWHQNLSLFYKDCF

Important features of the protein:**Signal peptide:**

amino acids 1-22

N-glycosylation site.

amino acids 294-298

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-34

Tyrosine kinase phosphorylation site.

amino acids 67-76

N-myristoylation sites.

amino acids 205-211, 225-231, 277-283

Amidation site.

amino acids 28-32

FIGURE 167

CCAGGACCAGGGCGCACCGGCTCAGCCTCTCACTTGTCTAGAGGCCGGGGAAGAGAAGCAAAGCGC
AACGGTGTGGTCCAAGCCGGGGCTTCTGCTTCGCCCTCTAGGACATACACGGGACCCCTAAGCTTC
AGTCCCCCAAACGCGCACCCCTCGAAGTCTTGAAGTCCAGCCCCGCACATCCACGCGCGGCACAGG
CGCGGCAGGCGGCAGGTCCCGGCCGAAGGCATGCGCGCAGGGGTCGGGCAGCTGGGCTCGGGC
GGCGGGAGTAGGGCCCCGCGAGGGAGGCGAGGCTGCATATTAGAGTTCGGGGCTGCGCCCTG
GGCAGAGGCGGCCCTCGCTCCACGCAACACCTGCTGCTGCCACCGCGCCGATGAGCGCGTGG
TCTCGCTGCTGCTGGGCGCGCGCTGCTCTGCGGCCACGGAGCCTTCTGCGGCCGCTGGTCAGC
GGCCAAAAGGTGTGTTTTGCTGACTTCAAGCATCCCTGCTACAAAATGGCCTACTTCCATGAAGT
GTCCAGCCGAGTGAGCTTTCAGGAGGCACGCTGGCTTGTGAGAGTGAGGGAGGAGTCCCTCCTCA
GCCTTGAGAATGAAGCAGAACAGAAGTTAATAGAGAGCATGTTGCAAAACCTGACAAAACCCGGG
ACAGGGATTTCTGATGGTGATTTCTGGATAGGGCTTGGAGGAATGGAGATGGGCAACATCTGG
TGCCTGCCCAGATCTCTACAGTGGTCTGATGGAAGCAATCCCAGTACCGAACTGGTACACAG
ATGAACCTTCCCTGCGGAAGTGAAAAGTGTGTTGTGATGTATCACCACCAACTGCCAATCCTGGC
CTTGGGGGTCCCTACCTTTACAGTGGAAATGATGACAGGTGTAACATGAAGCACAATTATATTTG
CAAGTATGAACCAGAGATTAATCCAACAGCCCCCTGTAGAAAAGCCTTATCTTACAAATCAACCAG
GAGACACCCATCAGAATGTGGTGTGTTACTGAAGCAGGTATAATTCCCAATCTAATTTATGTTGTT
ATACCAACAATACCCCTGCTCTTACTGATACTGGTGTGTTTGGAACTGTTGTTCCAGATGTA
GCATAAAGTAAGGAAGAACAATACTAGTCCAAACAGTCTACACTGTGGATTCAAAGAGTA
CCAGAAAAGAAAGTGGCATGGAAGTATAATACTCATTGACTTGGTCCAGAATTTGTAATTCT
GGATCTGTATAAGGAATGGCATCAGAACAAATAGCTTGGAAATGGCTTGAAATCACAAAGGATCTGC
AAGATGAAGTGAAGCTCCCCCTTGAGGCAATATTAAAGTAATTTTATATGTCTATTATTTC
TTTAAAGAAATATGCTGTGCTAATAATGGAGTGAGACATGCTTATTTTGCTAAAGGATGCACCCAA
ACTTCAAAGTCAAGCAATGAAATGGACAATGCAGATAAAGTTGTATCAACACGTCGGGAGTA
TGTGTGTAGAAAGCAATTCCTTTTATTTCTTTACCTTTTATAAGTTGTTATCTAGTCAATGTAA
TGTATATTGTATTGAAATTTACAGTGTGCAAAAGTATTTTACCTTTGCATAAGTGTGATAAAA
ATGAAGTGTCTAATATTTATTTTATGGCATCTCATTTTCAATACATGCTCTTTGATTAAAG
AACTTATTACTGTTGTCAACTGAATTCACACACACACAAATATAGTACCATAGAAAAAGTTTGT
TTTCTCGAAATAATTCATCTTTTCTGCTTTTGGTCAATGTCTAGGAAATCTCTTCAGA
AATAAGAGCTATTTTCAATAGTGTGATATAAACCTCCTCAAACATTTTACTTAGAGGCAAGGAT
TGTCTAATTTCAATTGTGCAAGACATGTGCCTTATAATTATTTTAGCTTAAATTAACAGATT
TTGTAATAATGTAACCTTGTAAATAGGTGCATAAACACTAATGCAGTCAATTTGAACAAAAGAA
TGACATACACAATATAAATCATATGTCTTACACGTTGCCTATATAATGAGAAGCAGCTCTCTGA
GGGTTCTGAAATCAATGTGGTCCCTCTCTTGCCCACTAAACAAAGATGGTTGTTCCGGGGTTTGGG
ATTGACACTGGAGGCAGATAGTTGCAAGTTAGTCTAAGGTTTCCCTAGCTGTATTTAGCCTCTG
ACTATATTAGTATACAAAGAGGTCATGTGGTTGAGACCAGGTGAATAGTCACTATCAGTGTGGAG
ACAAGCACAGCACACAGACATTTTAGGAAGGAAAGGAAGTACGAAATCGTGTGAAATGGGTTGG
AACCATCAGTGATCGCATATTCATTGATGAGGGTTGCTTGAGATAGAAAATGGTGGCTCCTTT
CTGTCTTATCTCCTAGTTTCTTCAATGCTTACGCCCTGTTCTTCTCAAGAGAAAGTTGTAACCT
CTGGTCTTCATATGTCCCTGTGCTCCTTTTAAACCAATAAAGAGTCTTGTCTTCTGGGGGAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 168

MSRVVSLLLGAALLCGHGAFCCRVSQKVCFADFKHPCYKMAYFHELSSRVSFQEARLACESE
GGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDGFWIGLWRNGDGQTSGACPDLYQWSDGSNSQ
YRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWDDRCNMKHNYICKYEPEINPTAPVEK
PYLTNQPGDTHQNVVVTEAGIIPNLIYVVIPTIPLLLLILVAFGTCCFQMLHKSkgRtKtSPNQ
STLWISKSTRKESGMEV

Important features of the protein:**Signal peptide:**

amino acids 1-21

Transmembrane domain:

amino acids 214-235

N-glycosylation sites.

amino acids 86-89, 255-258

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 266-269

N-myristoylation sites.

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-145, 212-217